

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2002, 19:49:46 : Search time 1914.27 Seconds
 (without alignments)
 12444.496 Million cell updates/sec

Title: US-09-993-874-1
 Perfect score: 1765
 Sequence: 1 tcagtgcaagactttaggga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table: IDENTITY_NUC
 Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

EST:*
 1: em_estba:*
 2: em_esthum:*
 3: em_estin:*
 4: em_estmu:*
 5: em_estov:*
 6: em_estpl:*
 7: em_estro:*
 8: em_htc:*
 9: gb_estl:*
 10: gb_est2:*
 11: gb_htc:*
 12: gb_gss:*
 13: em_gss_hum:*
 14: em_gss_inv:*
 15: em_gss_pln:*
 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	12.0	768	10	BM412281
2	188.8	10.7	721	10	BM413189
3	183	10.4	664	10	C83857
C 4	182.8	10.4	741	9	AU214129
5	180	10.2	882	10	BG445463
C 6	179	10.1	611	10	BJ129072
C 7	179	10.1	667	9	AI486799
C 8	178.2	10.1	604	9	AU219684
C 9	176.2	10.0	638	9	AU214473
10	175.8	10.0	582	9	AW621420
C 11	174.4	9.9	722	10	BJ139943
12	174.2	9.9	606	10	BI923203
13	174.2	9.9	651	10	C90519
14	171.2	9.7	612	10	BG130292
15	170	9.6	722	10	BM413019
16	169.8	9.6	713	10	BM406357
17	169.8	9.6	743	10	BM113456

18	168.6	9.6	569	9	AW218471
19	167.6	9.5	641	10	BG596946
20	167.4	9.5	522	9	AI488821
21	167.2	9.5	625	9	AW776080
22	167.2	9.5	631	10	BE435449
23	166	9.4	570	10	BG791252
24	165.6	9.4	640	10	BM110843
25	165.4	9.4	505	9	AW132800
26	164.8	9.3	529	10	BM143081
27	164.8	9.3	625	9	AI485586
C 28	163.8	9.3	558	9	AV680327
29	162.8	9.2	590	9	AW329421
30	162	9.2	1065	9	BE036528
31	161.2	9.1	520	10	BI269977
32	160.6	9.1	876	10	BG414540
33	160.4	9.1	593	10	BM157962
34	158.2	9.0	656	9	AW691106
C 35	156.6	8.9	694	9	AV860464
36	156.4	8.9	547	9	AV916926
37	156.2	8.8	535	9	AL389862
C 38	156.2	8.8	808	9	AW348381
39	155.6	8.8	644	10	BI309413
C 40	154.8	8.8	772	10	BE820931
41	153.2	8.7	618	9	AW776415
42	152.8	8.7	764	9	AA979869
43	152	8.6	497	9	BE020126
44	151.8	8.6	533	10	BE050505
C 45	151.4	8.6	683	10	BF492687

ALIGNMENTS

RESULT 1

BM412281

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM412281 768 bp mRNA linear EST 22-JAN-2002
 EST586608 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLEGS9P7 5' end, mRNA sequence.
 BM412281
 BM412281.1 GI:18263911
 EST.
 tomato.
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 768)
 Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai, J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 Unpublished (2002)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics Institute
 Seq primer: T3.
 Location/Qualifiers
 1. 768
 /organism="Lycopersicon esculentum"
 /cultivar="TA495"
 /db_xref="taxon:4081"
 /clone="CLEGS9P7"
 /clone_lib="tomato breaker fruit"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCUadapt; Site1: EcoRI; Site2: XhoI; supplier: Boyce Thompson Institute;

sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 219 a 143 c 195 g 211 t
ORIGIN

Query Match 12.0%; Score 211; DB 10; Length 768;
Best Local Similarity 56.3%; Pred. No. 2.4e-21;
Matches 425; Conservative 0; Mismatches 315; Indels 15; Gaps 1;
QY 782 ttgttagtagacttaaggttgtaagtgtgaagaaatttgagggcagcttttcttaaaa 841
DB 14 TTGAGATCGGAGCAGCAATTTGATATGATGAGAAATTCGACATTCCTGCTGGAG 73
QY 842 accatcacagaattacaaaatccccactattgtagtggccccctccagttatggtgttttg 901
DB 74 TTGATACAGAGCATAAAGTACGATTGGGCCATTGTACCGCCTATTGTTCTCGCCATT 133
QY 902 gctaaaaaccattagtgcatacaatacattttatagagcttaacggagaggttgctactgga 961
DB 134 GCTAAGAGCCCATATTGATGAATACATACATCTTTCATCAGTAAGAACAACAGTCATGTCGTG 193
QY 962 gtagctcttttagaaaagatgctgcagagcagtagcaaaagaggttgaaattacctgga 1021
DB 194 GCTGCACCATTTAGCGAAGCACTTGAAGATGCCGTCCGAGCCAAATTCCTTAACGCTAAA 253
QY 1022 atcatcacagatattggaattacgtgaagaaatttg-----ctgcgctgtaatg 1066
DB 254 CTCGGTCAGGGTTATGGAATACGGAAGCTGTGTCAGTGTGGCATGTTGTCATTT 313
QY 1067 attaccctcataatgctgtgtaaaacaggttcaactgaagagacccttccatacaattaaa 1126
DB 314 GCGAAGAAGAACCGTTTGATATTAAATTCAGGGGCGCATGTGTTACCGTTGTGAGGAACGACAG 373
QY 1127 gtaaaagtttagataacgctactgtggaaggcgctaggaccagggaggaagggcgaaata 1186
DB 374 ATGAAAATTGTCATCCGATACGGTTGCTCTCTGCCCGTAAACCAACCCGGTGAATTT 433
QY 1187 tgccttcaaaagtgaattgattgaaagattattacaacaatccggaagcaactattgat 1246
DB 434 TGCATTAGAGGTGATCAAAATCAGAAAGGTACTTGAATGACCCCTGAAGCTACAGCTAGA 493
QY 1247 actattgacaagatggtgtgcttcatctgagatattgaattatgcagcagcaagatgga 1306
DB 494 ACAATAGAAAAGAGGATGTTACACACTGGCGATATTGGATATATTGACGATCATGAT 553
QY 1307 aattcttttagttgctgattgaaagaacttataatacaaggggatatacaggttgcg 1366
DB 554 GAGCTTTTCATCGTCGATCGATTGAAGGAATTGATCAAAATCAAAAGGATTTCAAGTGGCG 613
QY 1367 cctgctgaactggaaaaactcgtcttttacaacaatccaaagtattgctgagcggtgtact 1426
DB 614 CTTGCTGAACTCGAAGCACTCTCTTGTCAACACCCCTAACATCTTCATCTGCTGCTGTC 673
QY 1427 ggaagctccgacgaattggtggacaattacactgctgctgtgtgtgtgttagaactcggc 1486
DB 674 CCATGAAGATGAACAGCGGGAGAGTTCAGTGGGTTTGTGTTGTAGATCAAAATGA 733
QY 1487 aagacgtgactgaaaggaaggttcaagattttat 1521
DB 734 TCTACCATTAAGATGAAGTGAAGGATTCAT 768

RESULT 2
BM413189 721 bp mRNA linear EST 22-JAN-2002
LOCUS BM413189
DEFINITION EST587516 tomato breaker fruit Lycopersicon esculentum cDNA clone
cLEG62P7 5' end, mRNA sequence.
ACCESSION BM413189
VERSION BM413189.1 GI:18264819

KEYWORDS EST.
SOURCE Lycopersicon esculentum
ORGANISM tomato.
REFERENCE 1 (bases 1 to 721)
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai,J., Bougri,O., Kirkness,E., Utterback,F., Van Aken,S., Rohnig,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics Institute
Seq primer: T3.
FEATURES
source
1..721
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG62P7"
/clone_lib="tomato breaker fruit"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="vector: pBluescriptSKMCladep; Site.1: EcoRI; Site.2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT 221 a 134 c 166 g 200 t
ORIGIN

Query Match 10.7%; Score 188.8; DB 10; Length 721;
Best Local Similarity 57.1%; Pred. No. 3.8e-18;
Matches 387; Conservative 0; Mismatches 282; Indels 9; Gaps 2;
QY 967 tcccttagaanaagatgctgcagaaagcagtagcaaaagaggttgaaattacctgggaatcat 1026
DB 8 TCACCTTGGGAAAGATGTTATGCAAGCATTTGCCACAAGTTTCCAAACATCATATTATT 67
QY 1027 caaaggatattgatttaactgaaactgctgcgctgtaatgattaccctcataatgctgt 1086
DB 68 TCAAGGATATGGACTACTGAACAAACCGGACGACGATTTCCGTCGTGCAACACTGAAGA 127
QY 1087 ga-----aaacaggttcaactggaagacccttggccatcacattaaagctaaagttttaga 1140
DB 128 AATGCTTCATCAGGGTTCTGTAGGAAGGTTAATGSCAAATTTCTGAAGCAAAATTTGTAGA 187
QY 1141 taacyctactctgggaaggcgttaggaccaggaagagaggaataatgctttcaaaagga 1200
DB 188 TCCGACACTGGAATGTCTGTCTGTCTGTCGTGAGAGGAGGCTCTCGATTAAGAGTCC 247
QY 1201 aatgattatgaaggatattacaacaatccggaagcaactattgatactattgacaaga 1260
DB 248 AATCATTAATGCAAGGTTACATCGGTGATCCCAAGAACACTTCTGTAACCTCTGATGCCATG 307
QY 1261 tgggttgcttcatctcggagatattggatattacgacgaagatggaaattcttttatagt 1320
DB 308 TGGGTGGTTGAGGACTGGTGATCTTTGTATTATATTGATCATCATGATACCTTTTGTGCTG 367
QY 1321 tgatgattgaaagaacttattaaatacaagggatatacaggttgccgctgctgtaactgga 1380
DB 368 AGACAGGCTGAAGGAGCTGATCAAAATATAGGGATACCAGGTGCTCTGCTGCTGAACACTAGA 427

```

Qy 1381 aaactgctcttacaacatcccaagattgctgtagcggtgttactgaggttcccgagca 1440
Db 428 ACAACTTCTCAATCTCACCTGCAATAGTAGATGCTGCTGTCAATACCATATCTGTGTA 487
Qy 1441 atttggtgacaattacactgctgctgtgtgtgtgttagaattctggaagcgtgactga 1500
Db 488 AGAAGCTGGTCAACTGCCATGGCGGTTGTGTGAGAGCTCCCAAGCACTCTGTGACAA 547
Qy 1501 aaaggaagttcaagattttatgtagcagcaagtcactcccaaaagcattcttcgagcgg 1560
Db 548 AGAACAAGTAGTATGATTTATTTCAAAACAGGTTGCTCCATATAAGAAAGAT---AAGCG 604
Qy 1561 tgcgtatttttagacagtatttcgaagagccctactgtaaaactcatcagaagaggagct 1620
Db 605 GTGGCATTTGTAGTCCATACCAAAAGTCCATCAGGGAAGATATTGAGAAGAGAAT 664
Qy 1621 ccgagaatatttgccca 1638
Db 665 AAAGGATTATTTACCA 682

RESULT 3
LOCUS C83857 664 bp mRNA linear EST 28-APR-1999
DEFINITION C83857 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSA391, mRNA sequence.
ACCESSION C83857
VERSION C83857.1 GI:2706789
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE 1 (bases 1 to 664)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pl,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source
1..664
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSA391"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

Query Match 10.4%; Score 183; DB 10; Length 664;
Best Local Similarity 55.6%; Pred. No. 2.7e-17;
Matches 351; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

Qy 1027 acaaggatgattgaactgaaactgctgctgctgaatgattaccctcctaatactgct 1086
Db 1 ACAAGTTATGGTGTACCGAAGCTTAGTCCAGCCTGTTTGTATATCCAACTGGATTAGT 60

Qy 1087 gaaaacaggttcaactggaagaccccttgcctacattataagctaaagttttagataaacgc 1146
Db 61 TAAATCTGTTCCGCTGGTATCCTCTTACCAATCAATGCTAAATCATCTCTCCAGA 120

```

```

Qy 1147 tactgggaagcgctaggaccagagaaagaggaataatgctttccttcaaaagtgaatgat 1206
Db 121 GACGTGTCAGAAATTTAGGTGAGTGAAGAGGGTGAATATCTGTATTAAGAGTCCAAATGT 180
Qy 1207 tatgaagaagatattacaacaatccggaagcaactattgatactattgacaagaatggttg 1266
Db 181 TATGTTGGTTATTATACAAATGAAGGCAACCAATCAAGTTCATAGATAAAGATGGATT 240
Qy 1267 gcttcattctggagatattgatalattacgagaaagatggaattcttcttatagtgatcg 1326
Db 241 CTTAAAAAAGTGTGATATTGGTTAGTGTGATGAAGATGGTTATTACTTTATCTGTCGATAG 300
Qy 1327 attaaagaacatttataacaaggaatatacagtttcgctgctgactggaactggaataatc 1386
Db 301 ATCAAGGAATTTGATCAAAATGTAAAGGTTTCCAACTACTCTCTGCCGAATTTGGAGGCATT 360
Qy 1387 gcttttacaacatcccaagattgctgtagcggtgttactgaggttcgaggtccggaagattgg 1446
Db 361 ACTATTATCTCATCCAAAAGTTGCAGATGCCGTGTGTAGTAGTCTTTCAAAAAGGTGATAT 420
Qy 1447 tggacaattacactgctgtgtgtgtgttagaattctggaagacgctgactgaaaaagga 1506
Db 421 GGGTGAAGTACCAAGAGGTTTGGTTGTTATTAAACAAAATGAATCTCTAACCGGAAAAAGA 480
Qy 1507 agttcaagattttattgagcacacagtcactcccaacaagcatcttcgagcggtgtcgt 1566
Db 481 ACTCTTAGATTGGGCCCCATCCAAAATTTGCAAAATTTATAAACAATTTACGTAATAAATCTTTAAGA 540
Qy 1567 attgtgacagatttcggaagagccctactggaactcatcagaagagagctccgaga 1626
Db 541 TTTCTATACCTGCAATTCCTAAATCTGCAACCGGTAAACTATTACGTAAATAATCTTTAAGA 600
Qy 1627 aatatttcccagcgagccacaaaatacaaaa 1657
Db 601 TTTTAATCCTCAAAAATATAAAAATAAANA 631

RESULT 4
LOCUS AU214129/c 741 bp mRNA linear EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone yk809f05 3', mRNA sequence.
ACCESSION AU214129
VERSION AU214129.1 GI:14852286
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 741)
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..741
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk809f05"
/clone_lib="unpublished oligo-capped cDNA library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others

```


[illegible]

```
Db 181 CAAGGTTACGGAATGACAGACGCGGACACAGTGTGGTGTATGTGCTTGGCAATTTGCTAAA 240
Qy 1073 cctcataatgctgtgaacaacaggttcaactgtgaagacccttgccatacatataaagctaaa 1132
Db 241 GAACCCCTTCGAATTAATCTGGACCATGTGGACTGTGTGTAGAAATGCTCAGATGAAA 300
Qy 1133 gttttagataacgctactgtggaagcgctgagaccagagagaagagcggaatatgcttt 1192
Db 301 ATTGTGGATCTGATACCTGCGCAATCTCTCCCTAGGAACCAATCTGGAGAGATTTGTATA 360
Qy 1193 caaagtgaatgattatgaaggatattacaacataataacacccgggaagcaactatgatactatt 1252
Db 361 AGAGCGCATCAAAATATGAAGGTTTACCTAAATGATCCAGAGGCCACTACCGGAACAATA 420
Qy 1253 gacaaagatggtgtgctcatcttgagatattgatatattacgacgaagatggaaatttc 1312
Db 421 GACAAAGAAAGGTGGTTACATACGGCGCATATTGGTTATTTGACATCGATGATGAGCTT 480
Qy 1313 tttatagttgactgattganaaactattataatacaaggagatattgctgctgctgct 1372
Db 481 TTCAATGTGGATCGTTTAAAGGAATGATAAAATACAAAGGATTTCAAGTGGCTCCTGCT 540
Qy 1373 gaactggaaaactgctttttacaacacccaagattgctgctgctgctgctgctgctgctgct 1432
Db 541 GAACCTGGAAGCCCTTCTCCTCAATCATCCCAATATTTTCAGATGCTGCTGTTTCAATG 600
Qy 1433 cggagcaaatgttggaacattacactgctgctgctgctgctgctgctgctgctgctgctgct 1492
Db 601 AAAGAGGACGAGCAGGAGAGACTTCTCTGCGCTTTTGTGTTCAGATCAAAATGGCTCCACA 660
Qy 1493 ctgactg 1499
Db 661 ATTACTG 667

RESULT 8
AU219684/c 604 bp mRNA linear EST 17-JUL-2001
LOCUS AU219684
DEFINITION AU219684 unpublished oligo-capped cDNA library, stage L1
Caenorhabditis elegans cDNA clone yk880b12 3', mRNA sequence.
ACCESSION AU219684
VERSION AU219684.1 GI:14857841
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 604)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykoha@lab.nig.ac.jp.
Location/Qualifiers
1. 604
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk880b12"
/clone_lib="unpublished oligo-capped cDNA library, stage
L1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 177 a 129 c 107 g 190 t 1 others
ORIGIN
```

```
Query Match 10.1%; Score 178.2; DB 9; Length 604;
Best Local Similarity 61.7%; Pred. No. 1.4e-16;
Matches 320; Conservative 0; Mismatches 193; Indels 6; Gaps 2;

Qy 1094 ggttcaactggagaccccttgcatacatataaagctaaagttttagataacgctactggg 1153
Db 591 GGATCCGTTGGAAAACCTTGCCTTCAAAATTTGGTTATGAAGATTTGTGGAACCAAGGACCGAC 532
Qy 1154 aaggcgctaggaccaggaagagaggaataatgctttcaagtgaatgataatgataa 1213
Db 591 AGAGAGCAACCGGTTAACCAACGAGGAGAAATTTGTGTGCTGCTCCAAACAAATATATGCTT 472
Qy 1214 ggatattacaacaatccggaagcaactattgatactattgscacaagaagggttggctcat 1273
Db 471 GGGTATCTCGGAAGACGAGAACCAAC---AGCCAGCACTGTCTCATCGATGGCTTCAC 415
Qy 1274 tctggagatatgatatattacgacgaagatggaaaattctcttatagttcgactgaaa 1333
Db 414 ACTGGAGACATTTGGTTATCTCAATGAGATGGAATTTGTTTATTTGTGTGATCGACTTAA 355
Qy 1394 gaacttataacaagggatatacaggttcgctgctgacagcggaatactgctttta 1393
Db 354 GAGCTAATCAAGGTCAAAAGGGCTTCAAGTACCACGCTGAACCTTGAAGATCTTCTTCTTA 295
Qy 1394 caacatcaagattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1453
Db 294 TCTCATCAAAAGATTAGAGATTGTGCAGTTTATGGAAATCCCAGATGCTAAAGCTGGAGAA 235
Qy 1454 ttacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1513
Db 234 TTACCTAAAGCATTTGTTG---TAAGAGCTGATAACACTCTCACTGGAACAAAGTTAAG 178
Qy 1514 gattttattgacacaaagtcactccaacaagcatcttcgagcggtgctgatttga 1573
Db 177 GACTTTGTTAAACCAAAAGTGTCTCTTACAAACAACTCGAAGGTGGTGTGCAATTTAT 118
Qy 1574 gacagtatccgaagggccctactggaatactcataga 1612
Db 117 GAAGAAATCCAAAGTCCGCGACTGGAAAGATTTTGACA 79

RESULT 9
AU214473/c 638 bp mRNA linear EST 17-JUL-2001
LOCUS AU214473
DEFINITION AU214473 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone yk813e08 3', mRNA sequence.
ACCESSION AU214473
VERSION AU214473.1 GI:14852630
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 638)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykoha@lab.nig.ac.jp.
Location/Qualifiers
1. 638
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk813e08"
/clone_lib="unpublished oligo-capped cDNA library, stage
```


Caenorhabditis elegans cDNA clone yk1152e02 3', mRNA sequence.
BJ139943
VERSION BJ139943.1 GI:18300109
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS 1 (bases 1 to 722)
Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1. 722
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1152e02"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 195 a 161 c 125 g 239 t 2 others
ORIGIN

Query Match 9.9%; Score 174.4; DB 10; Length 722;
Best Local Similarity 59.5%; Pred. No. 4.4e-16;
Matches 388; Conservative 0; Mismatches 251; Indels 13; Gaps 5;
Qy 957 ctggaggagctctttagaagaagctgcgagagcagtagcgaagaggt---tgaagt 1013
Db 698 CGGGAGCAGCTCCAGCTGGAAGGATTTGATTGAAGATTTGAAGAGGAGATACAAATT 639
Qy 1014 tacctggatcatcaagaagatgattgattaaactgaaactgtcgtctgtaagtattacc 1073
Db 638 TGAATATACATCAACAAGGATATGGAATGACGGAATGTTCAATGCTAGTCATCTCTCG 579
Qy 1074 tcataatgctgtgaaaca---ggttcaactggaagacccttgccatcacataaagcta 1130
Db 578 ATTTAAGAAATGATCAACCATATGGATCGTTGGAAACCTTGCTTCAATTTGGTTATGA 519
Qy 1131 aagtttttagataacgctactggaagcgctaggaccaggagaaagagcgaataatgct 1190
Db 518 AGATTGTGGACCAAGCAACAGACAGAGCAACCGGTTAAACAGAGGAGGAGAAATTTGG 459
Qy 1191 ttcaagtgaatgattatgaagaagattatcaacaatccggaagcaactattgtactata 1250
Db 458 TTGCGGTCCAAACAATATTGCTTGGGTATCTCGGAAGACCAGAACCA---GCCAGCA 402
Qy 1251 tgacaaagatggttggtggttcatttcattgagatattggtattatgaagcaagatgaaatt 1310
Db 401 CTGTCATCGATGGATGGCTTCACACTGGAGACATTTGGTTATCTCAATGAGGATGGAAT 342
Qy 1311 tctttatggttgctgattgaaagaacttataatacaagaagatcagattgctgctg 1370
Db 341 TGTATTATTGATGACCTTAAAGAGCTTAATCAAGTCAAGGGGTTCAAGTACCACAG 282
Qy 1371 ctgaactggaataatgcttttacaacatccaagattgctgctgctgctgctgctgctgag 1430
Db 281 CTGAACCTTGAAGATCTTCTCTATCTCATCCAAAGATTAGAGATTGTGCAGTTATTGAA 222
Qy 1431 ttccgagcgaatttgggacaaattacactgctgctgctgctgctgctgctgctgctgctg 1490

Db 221 TCCAGAGTGTAAAGCTGGAGAAATTACCTAAAGCAATTTGTTG---TAAGAGCTGATAACA 165
Qy 1491 cgtgactgaaaagaagtccaagatttttatgcagcagaagctactccaacaaagcattc 1550
Db 164 CTCTCAGTGAACAAGAGTTAAGGACTTTGTTAAACCAAGAGTGTCTCTTTACAACAAC 105
Qy 1551 ttccgagcggtgctgattttgtagacagattttccgaaagcgccctactctgaaa 1602
Db 104 TCGAAGGTGGTCTCGCAATTTAT-GAAGAAATTCCAAGTCCGCAGCTGGAAA 54
RESULT 12
BI923203 606 bp mRNA linear EST 18-OCT-2001
LOCUS EST543107 tomato callus Lycopersicon esculentum cDNA clone
DEFINITION cLEC79H20 5' end, mRNA sequence.
ACCESSION BI923203
VERSION BI923203.1 GI:16221846
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 606)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S. A., Teel
, J., Utterback, T., Van Aken, S., Ronning, C. M., Fraser, C. M., Martin
, G. B., Tanksley, S. D. and Giovannoni, J. J.
Generation of ESTs from tomato callus tissue (2001)
Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
source
1. .606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC79H20"
/clone_lib="tomato callus"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Boyce Thompson Institute; sequencing: The
Institute for Genomic Research; cLEC - Cotyledons of
seedlings 7-10 days post-germination were excised, cut at
both ends and placed on MS medium with no selection. Mixed
callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"
BASE COUNT 207 a 84 c 138 g 177 t
ORIGIN
Query Match 9.9%; Score 174.2; DB 10; Length 606;
Best Local Similarity 57.8%; Pred. No. 5.1e-16;
Matches 339; Conservative 0; Mismatches 233; Indels 15; Gaps 1;
Qy 810 tgaagaaatttgaggcgcacttttcttaaaacacatcacagaaattacaaaatccccacta 869
Db 16 TCAAGAAATTTGATGTTGTTACATGCTTGGAAATTCGTACAAAGTACAAAGTCACAATTG 75
Qy 870 ttgtagtgccctccagttatggtgttttggctaaaagcccatagtcgatcataacg 929
Db 76 GACCTTTGTACCACCTATATTTTGGCTATTGCTTAAAGTCAAGTACTGATATATG 135
Qy 930 attatcagagcttaacggaagtgtgctactgaggagctctctttagaaaaagatgtcgag 989

```

Db 136 ACCTTTCATCAGTAGGATAGTCATTTCTGGGGCAGCACCATTAGGAAGAAGCTTGAAG 195
Qy 990 aagcagtagcaagagggttgaaattacctggaatcatacagaaggaatggatgaactga-- 1047
Db 196 AGGTAGTTGGAATTAATTTCTTAATGCTAAACTTGGTCAGGGTTATGGAATGACGGAGG 255
Qy 1048 -----aactgtgcgtgtaagtattaccctccataaataatcgtgtaaaaaaag 1094
Db 256 CCGGPGCACCTATAACAATAGCTTAGCATTTGCAAAAAGAACCTTTGAGATAAATCAG 315
Qy 1095 gtcaactgggaagacccttgccatacatataagctaaagctaaagtttttagataacgcactcggga 1154
Db 316 GGGCATGTGGCACACTATAAGGAATGCTGAGATGAATAATTTGTGATCCAGATAGTGGA 375
Qy 1155 aggcgttaggaccaggagaagagcgaaatagctttccaagtgaatgattatgaaga 1214
Db 376 TTTCTCTCTCGAATAAACC CGGAGAAATTTGCAATAGAGGTGACCAAAATTAATGAAG 435
Qy 1215 gatattacaacaatccggaagcaactatgatactatcattgacaaagaatggttgcctcatt 1274
Db 436 GTTATTAATTAATCCTCAGCGAACCTAAAGAACTAAGATTAAGAAAGGATGGTTACATA 495
Qy 1275 ctggagatattggaattacacagcagagatggaaattctttatagttgattgattgaaga 1334
Db 496 GTGGTGATATAGGCTATATTGACGATGATGATGACCTTTTATTGTTGGATCGATTAAAGG 555
Qy 1335 aacttattbaatacaggggatagcaggttgccgctgctggaactggaa 1381
Db 556 AATTGATAAATAAAGGATTTCAAGTAGCAGCACCTGCTGAGCTTCAA 602

RESULT 13
C90519 651 bp mRNA linear EST 20-APR-1998
LOCUS C90519 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION discoideum cDNA clone SSI661, mRNA sequence.
ACCESSION C90519
VERSION C90519.1 GI:3060139
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., P.I.M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE Developmental cDNA in Dictyostelium discoideum
JOURNAL Unpublished (1998)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp.
FEATURES
    source
        251 a 83 c 110 g 207 t
        /organism="Dictyostelium discoideum"
        /strain="AX4"
        /db_xref="taxon:44689"
        /clone="SSI661"
        /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
        /dev_stage="slug"
BASE COUNT 251 a 83 c 110 g 207 t
ORIGIN

Query Match 9.9%; Score 174.2; DB 10; Length 651;
Best Local Similarity 55.6%; Pred. NO. 4.9e-16;
Matches 355; Conservative 0; Mismatches 283; Indels 1; Gaps 1;
Qy 1127 gctaaagttagataaagcactactggaagcgcttagaccaggaagagcgcaataa 1186
Db 1 CCTAAATCATCTCTCTAGAGACTGCTGAGAAATTTAGGTATGGGTGAAGAGCGTGAATC 60

```

```

Qy 1187 tgccttcaagaatgaaatgattatgaagaagagatattacaacaatccgggaagcaactattgat 1246
Db 61 TGTATTAAAGGTCCAAATCTTATGTTGGTTATTATAACAATGAA-AGGCAACCAATGAA 119
Qy 1247 actattcaagaagtggtcgtctcattctcgagatattggatatctagacacgaagatgga 1306
Db 120 GTCATAGATAAAGATGGATTCCTTAAAACTGGTGATATTTGGTTATGTGATGAAGATGT 179
Qy 1307 aattctttatagttgatcgatgaagaacttataatacaagaaggatcaggttgcg 1366
Db 180 TATTTCTTTATCATCGATAGATCAAGGAATGATCAAAATGTAAGGTTTCCAAGTACCA 239
Qy 1367 cctgctgaactggaaaatcctgcttttaacaacatccaaagtattcgtgatcgcggtgttact 1426
Db 240 CCGTCTGAATTAGAAAGCATTTACTATTATCTCATCCAAAAGTTGCAGATCCCTGTGTAGTA 299
Qy 1427 ggaagtcctggcgaagaattgggtgacaattaccctgctgctgtgtgtgtgtagaactcgc 1486
Db 300 GGTCTTTCAAAAGGTGATGGGTGAAGTACCAAGAGGTTTCGTTGTTATTAACAATAAT 359
Qy 1487 aagacgctgactgaaaaggaagtccaagatttttattgcagcacagaagtcactccaacaag 1546
Db 360 GAATCTCTTACTGAAAAGAACTCTTTGGATTGGGCTCACCCAAAATTTGCAAAATTATAAA 419
Qy 1547 catcttcagggcgggtgctgattttagacagatattccgaaaggccctactactggaaaactc 1606
Db 420 CATTTTCAGAGGTGGTATCTTTTTTTCATCTGCAATTCCTTAAATCTGCAACTGCTAAACTA 479
Qy 1607 atcagaaggagcctcgagaataatttgcagcagcagcacaacaaatacgaatttaagtt 1666
Db 480 TTACGTAAATAATCTTAAAGATTTTAAATCTTAAATAATTTTAAATAATTTTAAAAAAGGTGT 539
Qy 1667 caatgattgctttagttctaaatgtataaacaagtttttagaacctaatcatatcat 1726
Db 540 TAAATTTCCACTTTCACCAACAAGATTTTCAAAATTAATTTATTTATATAATAGTAGGAAA 599
Qy 1727 tcaatatacaacaaaaaataaaaaaataaaaaaataaaaaa 1765
Db 600 CCAAAATGCCAACAAATAAAATTTTAAAGATATATCATATA 638

RESULT 14
BG130292
LOCUS EST475938 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION cTOF29J22 5' sequence, mRNA sequence.
ACCESSION BG130292
VERSION BG130292.1 GI:12630480
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
    source
        1..612
        /organism="Lycopersicon esculentum"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="cTOF29J22"
        /clone_lib="tomato shoot/meristem"

```

/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

BASE COUNT 189 a 104 c 137 g 182 t
ORIGIN

Query Match 9.7%; Score 171.2; DB 10; Length 612;
Best Local Similarity 58.8%; Pred. No. 1.4e-15;
Matches 315; Conservative 0; Mismatches 218; Indels 3; Gaps 1;
QY 1173 aaagaggcgaataatgcttcaaaagtgaatgattatgaaggatattacaacaatccgg 1232
DB 24 AACCGGTGAATTTGCATTAGAGTGATCAAAATCATGAAAGGTTACTGTGAATGACCCGTG 83
QY 1233 aagcaactatgatactattgacaagaatggttggtcttcattctggagatattgatatt 1292
DB 84 AAGCTACAGCTAGAACATAGAAAAGAGGATGCTTACACACTGGCGATATTGGATATA 143
QY 1293 acgacgaagatggaattcttatttagttgattgattgaaagaacttattataatacaagg 1352
DB 144 TTGACGATGATGATGAGCTTTTCATCGTGGATCGATTGAAGGAATTGATCAAAATACAAAG 203
QY 1353 gatatacaggttcgctgctgcaactggaaatctgctctttcaacatacaagattgctg 1412
DB 204 GATTTCAAGTGGCGCTGCTGAACFCGAAGCACCTTCTGTCAACCCACCTTAACATTTCTG 263
QY 1413 atgcgggtgctactgaggtccggacgaatttggggcaattacacctgctgctgtgttg 1472
DB 264 ATGCTGCTGTTCTCCCAATGAAGATGAACAGCGGAGAGATTCCAGTGGCTTTGTG 323
QY 1473 tgttgaatcbtggcaagcgtgctgcaaaaggaaagttcaagatttttattgacgacaaag 1532
DB 324 TTAGATCAATGGATCTACCATTAAGATGAAGGATGAAGGATTTTCATCTCCAAAGAGG 383
QY 1533 tcaatcccaaaagcattctcgagcggtgctgatttggtagacagtagtattccgaaggcc 1592
DB 384 TGATATTCTTAAGAAAT--AAAGCGTGATTTTTCGTGGAGCGGATCGCAAAATCTC 440
QY 1593 ctactggaaaactcatcagaaggagctccgagaataatttggccagcgagcaccataat 1652
DB 441 CATCAGGAAAAATTTCTGAGAAAGACTTAAGAGCTAGACTGCTGCTGTATTTCAAT 500
QY 1653 caaaattataagttcaatgattgctttagttcttaaaatgtatatacaaaagtttt 1708
DB 501 GATGGAGATACTATATGTTGTTGATTTGTACCCCTTATGTTTCGATCAACATAATTT 556

RESULT 15
BM413019 722 bp mRNA linear EST 22-JAN-2002
LOCUS EST587346 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION CLBG61P24 5' end, mRNA sequence.

ACCESSION BM413019
VERSION BM413019.1 GI:18264649
KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

REFERENCE 1 (bases 1 to 722)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S. A., Tsai
J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning
C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and Giovannoni, J.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
COMMENT Unpublished (2002)
Contact: CUGI

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..722
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG61P24"
/clone_lib="tomato breaker fruit"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKMvAdapt; Site_1: EcoRI;
Site_2: XhoI; Supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research;
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

BASE COUNT 216 a 123 c 163 g 220 t
ORIGIN

Query Match 9.6%; Score 170; DB 10; Length 722;
Best Local Similarity 58.2%; Pred. No. 1.9e-15;
Matches 318; Conservative 0; Mismatches 225; Indels 3; Gaps 1;
QY 1094 ggttcaactgaaagacccttgcacatactaaagctaaagtttagataaagcactactgg 1153
DB 37 GGTCTGTAGGAAGGTTAATGGCAANTTCTGAAGCAAAATTTAGATCCGACACTGGA 96
QY 1154 aaggcgctaggaccaggagaaggcgaaatctgcttccaaagtgaaatgattatgaaa 1213
DB 97 ATTGCTCTGTGTCCTGGTGAGCAAGGAGAGCTCTGGATTAAAGGTCCAATCATTTATGCAA 156
QY 1214 ggataattacaacatccgggaagcaactattgatactattgacaaagagtggttggcttcat 1273
DB 157 GGTACATCGGTGATCCGAAGAACACTTCTGAACCTCTGATGCCATGTGGTGGTTGAGG 216
QY 1274 tctgagatattggatattacgagagagatggaattcttcttatttagtcgactgaaa 1333
DB 217 ACTGCTGATCTTTGTTATATTGATCATCATGGATACCTTTTTCGTAGACAGGCTGAAG 276
QY 1334 gaacttattaaatacaagggatatacaggttgcgctgctgactggaactggaatactctgcttta 1393
DB 277 GAGCTGATCAAAATPAAGGGATACCAGGTTGCTCTGCTGAACTAGAACAACTTCTTCAA 336
QY 1394 caacatccaaagtattgctgagtcggggtgttactggaggttccggagcaatttgggtgacaa 1453
DB 337 TCTCACCTGAAATAGTAGTCTGCTCTCATACCATATCTCTGATGAAGAAGCTGGTCAA 396
QY 1454 ttactgctgctgttgttgtttagaatcttggcaagcgtgactggaagaagaagttcaa 1513
DB 397 CTGCCCCATGGCGTTTGTGTGAGAGCGTCCCCAAACACTCTTTGACAAAGACAAGTGATT 456
QY 1514 gattttattgagcacaagtcactcccaaaaagcaatcttcgaggcggtgtgtgtatttga 1573
DB 457 GATTTTATTTCAAAACAGGTTGCTCCATATAAGAAAGAT---AAGCGGGTGGCATTTGTT 513
QY 1574 gacagtattccgaaggcccttacttgaaaaactcatcagaagaaggagctccgagaaatatt 1633
DB 514 AGCTCCATACCAAAAAGTCCATCAGGGAAGATATTGAGAAGAGATTAATAAGATTTCATT 573
QY 1634 gcccaag 1639
DB 574 TACCAG 579

Search completed: September 7, 2002, 21:31:54
Job time: 6128 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2002, 20:03:21 ; Search time 77.82 Seconds
(without alignments)
5571.102 Million cell updates/sec

Title: US-09-993-874-1
Perfect score: 1765
Sequence: 1 tcagtcgaagacttttaggga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582	33.0	1659	1	US-08-231-729B-1
2	581.4	32.9	1659	1	US-08-231-729B-2
3	580.4	32.9	1725	2	US-09-380-051B-19
4	584.4	31.4	5620	2	US-08-793-170-21
5	584.4	31.4	5620	3	US-08-892-873-21
6	584.4	31.4	5620	4	US-09-334-765A-21
7	584.4	31.4	5620	4	US-09-356-575E-21
8	584.4	31.4	5620	4	US-09-333-820-21
9	582.8	31.3	1811	3	US-08-867-352-22
10	582.8	31.3	5789	3	US-08-862-431-32
11	582.8	31.3	5791	3	US-08-862-431-31
12	582.8	31.3	5793	3	US-08-862-431-29
13	582.8	31.3	5793	3	US-08-862-431-30
14	582.8	31.3	5818	2	US-08-536-559A-3
15	582.8	31.3	5819	2	US-08-536-559A-2
16	582.8	31.3	5819	3	US-08-862-431-27
17	582.8	31.3	5819	3	US-08-862-431-28
18	582.8	31.3	5938	2	US-08-536-559A-4
19	582.8	31.3	6092	2	US-08-536-559A-1
20	582.8	31.3	6092	3	US-08-862-431-26
21	582.8	31.3	11616	1	US-08-196-259-2
22	580.6	31.2	5427	3	US-09-282-996-2
23	550.2	31.2	1722	4	US-08-875-277A-1
24	550.2	31.2	1722	3	US-08-718-425-1
25	549.8	31.2	1650	1	US-08-354-240A-1
26	549.8	31.2	1722	4	US-09-380-051B-1
27	549.8	31.2	2445	1	US-08-122-520C-8

28	549.8	31.2	5405	3	US-09-282-996-1	Sequence 1, Appl
29	549.8	31.2	6044	1	US-08-316-950-18	Sequence 18, Appl
30	549.8	31.2	6044	5	PCT-US95-12642-18	Sequence 18, Appl
C 31	549.8	31.2	6047	1	US-08-316-950-12	Sequence 12, Appl
C 32	549.8	31.2	6047	5	PCT-US95-12642-12	Sequence 12, Appl
33	549.8	31.2	6171	1	US-08-316-950-17	Sequence 17, Appl
34	549.8	31.2	6171	5	PCT-US95-12642-17	Sequence 17, Appl
35	548.4	31.1	2009	4	US-09-380-061B-17	Sequence 17, Appl
36	547.8	31.0	1644	1	US-07-903-047-7	Sequence 7, Appl
37	547.8	31.0	1644	1	US-09-380-061B-15	Sequence 15, Appl
38	546.2	30.9	1644	3	US-09-111-752-13	Sequence 13, Appl
39	544.6	30.9	1908	1	US-08-460-934-8	Sequence 8, Appl
40	544.6	30.9	1908	2	US-08-782-118-8	Sequence 8, Appl
41	544.4	30.8	1704	1	US-08-460-934-5	Sequence 5, Appl
42	544.4	30.8	1704	2	US-08-782-118-5	Sequence 5, Appl
43	542.2	30.7	1656	3	US-09-111-752-8	Sequence 8, Appl
44	541.6	30.7	5157	2	US-08-474-169-7	Sequence 7, Appl
45	539	30.5	1656	3	US-09-111-752-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-231-729B-1
; Sequence 1, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shubei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIRELY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231.729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1659
; OTHER INFORMATION: "Xaa" at codon 409 is either Glu or Asp;
; OTHER INFORMATION: "Xaa" at codons 28, 32, 112, 130, 142, 190, 212, 217, 222,
; OTHER INFORMATION: 329, 336, 386, 436, 512, and 532 is either Arg, Ser or Gly

TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380.061B

FILING DATE: 25-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998

APPLICATION NUMBER: GB 9707468.8

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B. J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 124-725

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4000

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 1725 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 32..1675

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

us-09-380-061B-19

Query Match 32.9%; Score 580.4; DB 4; Length 1725;
Best Local Similarity 60.08; Pred. No. 8.2e-113;
Matches 1006; Conservative 0; Mismatches 661; Indels 9; Gaps 2;

Qy 29 gaagaagaaacattaggcatggagagcgtcctcgtgatagatccatcctcggtcgca 88
Db 38 GATGCAAAATATTATCCAGCGTCCAGCGCATTTCTATCTCTTTGGAGGATGGAACGTCT 97
Qy 89 ggacaacaattatccatcatctgtataaaattgcatactcttctcgtgaagaata----- 142
Db 98 GGAGAACAAATTCGCAAAAGCAATGAAGAGGTATGCACAGGTTCGAGGACAAATTCGCTTT 157
Qy 143 atcagatctacaaatgaagtagtaatacatgctcaaatatttgaaccagctgcgc 202
Db 158 ACTGATGCCCATGACAGGTAAATATTACATATTCGCGAATATTTTGAATGGCTTGGCGG 217
Qy 203 ttagctgttagtagaacaatatggcttgaaatgaacaacaatgttgggtgtatgcagt 262
Db 218 TTAGCGAAACTATGAGAGGTACGGACTTGGTTTACAACACACCATTTGCTGTGTAGC 277
Qy 263 gaaacaataaactttttaaactcctgtccttctgtccttcttatacttaggaataccagta 322
Db 278 GAAATTCCTCTAGTTTTTATGCTCTGATCGCGGGCTCTATTTATTTGGAGTTGGAGTT 337
Qy 323 gcaacataatgatgatcacagatggagagtaactggtcatttgaatatatacaaaa 382
Db 338 GCATCAACAATGATATTACAAAGAGTGAATTTACAAACAGTTTGTCCCATATACAA 397
Qy 383 ccaactatcgttttagttcaagaagaagcactcccgcttattctgagagtacagcaaat 442

Db 1478 GCCGTTCTGCTTTAGAGGAGGCAAAACGATGACTGNACAAAGTGTGATGATTATGTT 1537
Qy 1523 gcagcacaagtcactcccaacaaagcatcttcgagcgcggtgctgattttgtagacagttatt 1582
Db 1538 GCGGACAAAGTAACCTGCTTAAAGCGTTTACGTGGAGGAGTTAAGTTTGTGGACGAAGTA 1597
Qy 1583 ccgaaagccctactggaactcatcagaaagagagctccgagaaatatttggccacgaga 1642
Db 1598 CCTAAGGCTTAACCTGGAAGAGTTGATGGAAGAAATAATCAGGGAGATGCTTT--ATGATG 1654
Qy 1643 gcacccaaatcaaatataagttcaatgtattgtctttagttctcaaaatgtatata 1698
Db 1655 GGAAGAAATCCAAATGTAAATTCCTTCGGTTTACTATATATTCTTAACGAAATTTCTA 1710

RESULT 4

US-08-793-170-21
; Sequence 21, Application US/08793170
; Patent No. 5994128
; GENERAL INFORMATION:
; APPLICANT: FALLAUX et al.
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT
; TITLE OF INVENTION: ADENOVIRUS TO BE USED IN GENE THERAPY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
; STREET: 260 SHERIDAN AVENUE, PO BOX 60039
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08793170
; FILING DATE: 25-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/00326
; FILING DATE: 14-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201728.3
; FILING DATE: 26-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201611.1
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER, BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: INGE.002.000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)328-4400
; TELEFAX: (650)328-3377
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "plasmid"
US-08-793-170-21

Query Match 31.4%; Score 554.4; DB 2; Length 5620;
Best Local Similarity 59.2%; Pred. No. 2.7e-107;
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;
Qy 2 cagtcgaagacttttaggatcaaatggaagaagaaacattagcgatggagcgtcct 61

Db 1265 CATTCCGGTACTGTGTGTAATAAATGGAAGACGCCAAAACATAAAGAAAGGCCCGCGGCCA 1324
Qy 62 cgtgatatagccctccctcgtcgcgcaggaacaacaattatataccaatcattgtataaaatt 121
Db 1325 TTCTATCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAGAAATAC 1384
Qy 122 gcatttttctcctgaagca-----taategatcgtcgtacatacaaatgaagtaatatat 175
Db 1385 CCCCTGGTTCCTGCAACAATTTGCTTTACAGATGCACATATCGAGGTGAACATACGCTAC 1444
Qy 176 gctcaaatatttgaaccagctgcgcttagcttagttagtataagaacaatatggttgaat 235
Db 1445 GCGGAATACTTTCGAAATCTCCGTTTCGGTTGCGAGAAGCTATGAACAGATATGGGCTGAAT 1504
Qy 236 gaaacaatgttgtgggtgtgtagcagtgaaaaacaataaaaacttttttaactcctgtcctt 295
Db 1505 ACAAATCACAGATCGTCGTATGAGTGAATAAATCTCTCTCAATTTCTTTATCCGGGTGTTG 1564
Qy 296 gctgctttatacttaggaataccagtagcaacatacaaatgatatglacacagatggagag 355
Db 1565 GCGCGCTTATTTATCGGAGTTTCAGTTTCGCCCGCGGAACGACATTTATATGAACGTGAA 1624
Qy 356 ttaactggtcatttgaatatatacaaaacaaactatcatgttttagttcgaagaagacactc 415
Db 1625 TTGCTCAACAGTATGAACATTTTCGAGCTTACCGTAGTGTGTTGTTTCCAAAAGGGTGTG 1684
Qy 416 ccgcttattctgagtagcacgcaaaatcgaatttcataaaaaagtcgtagttatcagat 475
Db 1685 CAAAAAATTTTGAACGTGCAAAAAAATTTACCAATATCCAGAAAAATTTATATCATGAT 1744
Qy 476 agcatgtacgacattaatgcggttgaaatgcgtatctatctactttgttgacgttatactgac 535
Db 1745 TCTAAAACGGATTTACCGGATTTTCAGTCGATGTACACGTTTCGTCACATCTCATCTACT 1804
Qy 536 cacaccttgcattgcatttacacaaaagatttgccttgccttgccttgccttgccttgccttgc 595
Db 1805 CCCGTTTTAATGAATACGATTTTGTACCAGAGTCTTGTATCGTGACAAACAAATTCGA 1864
Qy 596 ttaattatgcattcattgcgaactggttgcctgaaggtgtagtagtagtagtagtagtagtag 655
Db 1865 CTGATAATGAATTCCTCTGGATCTACTGGTTTACCTAAGGGTGTGGCCCTTCCGCATAGA 1924
Qy 656 agtctaactataagattcgttcgaatagcaggaatccactttatggaactcgtacaggttcca 715
Db 1925 ACTGCTGCGTCAGATTTCTCGCATGCCAGAGATCTATTTTGGCAATCAAAATCATCCG 1984
Qy 716 caaacatcaattcttcttagtaccgttccatgccttgccttgccttgccttgccttgccttgc 775
Db 1985 GATCTGCGATTTTAAAGTGTGTTCCATTCATACGGTTTGGAAATGTTTACTACATCTC 2044
Qy 776 tcttactttgttagtaggacttaaggttgaatgttgaagaataattgagggcgacattttc 835
Db 2045 GGATATTTGATATGTGGATTTTCGAGTCTCTTAATGTATAGATTTTGAAGAGAGCTGTTT 2104
Qy 836 ttaaaaccatacacgaattacaaaatccccactatgttagtggccccctccagttatggtg 895
Db 2105 TTACGATCCCTTCAGGATTAACAAATTTCAAGTGGTGTCTAGTAGTACCAACCGTATTTTCA 2164
Qy 896 ttttggctaaaagccatttagtcgatcaatcacgatttalcgagcttaacggaagtgtct 955
Db 2165 TTCTTCGCCAAAGCACCTCTGATTGACAAATACGATTTATCTAATTTACAGAAATTTGCT 2224
Qy 956 actggaggagctccttttaggaaaaagatgcgcagagcagtagcaagaaggttgaatta 1015
Db 2225 TCTGGGGCGCACCTCTTTTCGAAAGAACTCGGGGAAGCGTTGCAAAACGCTTCCATCTT 2284
Qy 1016 cctggaatcatacaaggatattggaataactgaactgctgcgtgtaattgattaccct 1075
Db 2285 CCAGGGATACCAAGGATATGGGCTCAGTACAGATACATCAGCTATTCTGATTACACCC 2344
Qy 1076 cataatgctgtgaaacaggttcaactggaagaccttgcacatacatataaagtaagt 1135

```
Db 2345 GAGGGGATGATAAACCGGGCGCGTGGTAAAGTGTGTCCATTTTGAAGCGAAGGTT 2404
Qy 1136 ttgataacgctactgggaagcgctagaccaggagaagaagcggaataatgcttcaa 1195
Db 2405 GTGGATCTGGATACCGGAAACGCTGGCGTTAATCAGAGAGCGGAATTATGTCTCAGA 2464
Qy 1196 agtgaatgattatgaagggattatacaacaatccggaagcaaatattgatactattgac 1255
Db 2465 GGACCTATGATTATGTCGGGTATGTAAACAATCCGGAAGCGACCAACGCTTGTGATGAC 2524
Qy 1256 aaagatggttggttccttcctggagatatgtgatatcagcaggaagatggaaattctttt 1315
Db 2525 AAGGATGATGGCTACATCTGGAGACATCTTACTGGGACGAAGACAACTTCTTC 2584
Qy 1316 atagtgtatgatgaagaacttatataacaaaggaatcaggttgccctgctgaa 1375
Db 2585 ATAGTTGACCGCTTGAAGCTCTTAATTAATACAAAGGATATCAGGTGGCCCCCGCTGAA 2644
Qy 1376 ctgaaaaatcgtcttttacaacatccaagttatgctgatcggggtgttactggagttccg 1435
Db 2645 TTGGAATCGATATTGTTACAACACCCCAACATCTTCGACGCGGGGTGGCAGGTCTTCCC 2704
Qy 1436 gacgaatttgtagacaattaccctgctgctgtgtgtgttagaatcttggcaagacgctg 1495
Db 2705 GACCATGACGCGCGGTGAACCTTCCCGCCCGCGTGTGTGTTTGGAGCACGGAAGACGATG 2764
Qy 1496 actgaagaggaagttcaagattttatgacgacacaagtcactccaacaagcatcttoga 1555
Db 2765 ACGGAAAAGAGATCGTGGATTACGTGCCAGTACCGAAGCTCTTACCGGAAACTCGACGCRAGA 2824
Qy 1556 ggcgggtgctgattgttagacagttatccgaaagccctcaggaactcagaaactcatcagaag 1615
Db 2825 GGAGGAGTGTGTTGTGGACGAAGTACCGAAGCTCTTACCGGAAACTCGACGCRAGA 2884
Qy 1616 gagctccagagaat 1629
Db 2885 AAAATCAGAGAGAT 2898

RESULT 5
; Sequence 21, Application US/08892873
; Patent No. 6033908
; GENERAL INFORMATION:
; APPLICANT: FALLAUX et al.
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT
; TITLE OF INVENTION: ADENOVIRUS TO BE USED IN GENE THERAPY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: RAE-VENTER LAW GROUP, P.C.
; STREET: 260 SHERIDAN AVENUE, PO BOX 60039
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892.873
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/793.170
; FILING DATE: 25-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/00326
; FILING DATE: 14-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201728.3
; FILING DATE: 26-JUN-1995
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201611.1
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER, BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: INGE.002.000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)328-4400
; TELEFAX: (650)328-3377
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "plasmid"
; US-08-892-873-21
```

```
Query Match 31.4% Score 554.4; DB 3; Length 5620;
Best Local Similarity 59.2%; Pred. No. 2.7e-107;
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;
```

```
Qy 2 cagtgcaagactttaggagatcaaaatggagaagaacaacattagcattggagagcgctct 61
Db 1265 CATTCGGGTACTGTGTGTAATAATGGAGAGCGCCAAACAATAAAGAAAGGCCCGCGCA 1324
Qy 62 cgtgatatagtcctatcctgctgcgagggacaacaattataaccaatcattgtataaaatt 121
Db 1325 TTCTATCTCTAGAGGATGGAAACCGCTGGAGAGCAACTGCATAGGCTATGAAGAAATAC 1384
Qy 122 gcatctttctcgaagcaaa-----taatogatgctcacaacaaatgaagtaataatcat 175
Db 1385 GCCCTGGTTCCTGGAAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTAC 1444
Qy 176 gctcaaatatttgaaccagctgcgcttagctgttagtatagacaacatacgcttgaat 235
Db 1445 CCGGAATACTTCGAAATGTCGCTGTCGGTGGCAGAACGATATGAACAGATATGGCGCTGANT 1504
Qy 236 gaaacaatggttggttggtatgtagcagtgaaacaacataaaacttttttaaccctgtctct 295
Db 1505 ACAAATCACAGAAATCGTGTATGTCAGTGAAACATCTCTTCAATTCTTTATCCCGGTGTG 1564
Qy 296 gctgctttatctacttaggaatcccgtaggaacatacaaatgatatgacacagatggagag 355
Db 1565 GCGCGTTTATTTATCGGAGTTGTCAGTTTCGCGCCGCGAACGACATTTATTAATGAACGTGA 1624
Qy 356 ttaactggtcatttgaatatatacaaaacaaactatcattgtttagtccaaagaaagcactc 415
Db 1625 TTGCTCAACAGATATGAACATTTCCGACGCTACCGCTAGTGTGTTTGTTCAAAAGGGGTG 1684
Qy 416 ccgcttattctgagatcacagcaaaaactaagtttcttaaaaaagtcgtagtattcgat 475
Db 1685 CAAAAATTTTGAAGCTGCAAAAAAAATTAACCAATAATCCAGAAAAATTAATATCATGGAT 1744
Qy 476 agcatgtacgacatataatggttggaatgctatactacacttctgacagcttactgac 535
Db 1745 TCTAAAACGGATTACCAAGGATTTTCAGTCGATGTACAGTTTCGTGCATCTCATCTACCT 1804
Qy 536 cacacctttagatccattgtacacaaaagattttagtcccccttgaaaaaatcgca 595
Db 1805 CCCGTTTTAATGATAGGATTTTGTACCAGAGTCTTTGATCGTGACAAAACAATTTGCA 1864
Qy 596 ttaattatgtcatcatcttggaacaactggaattgctgaaggtgtagtagcagccataga 655
Db 1865 CTGATATGAATTCCTCTGGATCTACTGGGTACTAAGGCTGTGGCCCTTCCCGCATAGA 1924
Qy 656 agtctaactataagattcgtttcataggaaggaatcccatattatggcactcgtacagttcca 715
Db 1925 ACTGCCCTGCGTCAGATTTCTCGCATGCCAGAGATCCTATTTTGGCAATCAAAATCATTCG 1984
```

Qy	716	caaacatcaattcttctcttagtacggttcccatcatgaccttggaaatgtttactacatta	775
Db	1985	GATAC ¹ TCCGAT ¹ TTAAGT ¹ CTG ¹ TCCAT ¹ TCCAT ¹ CACG ¹ TT ¹ TGGAAT ¹ GT ¹ TACT ¹ TACAC ¹ TC	2044
Qy	776	tcttaccttgtagtaggaccttaaggtgttaatggttgaagaaattttagggcggaacttttc	835
Db	2045	GGATAT ¹ TTGATAT ¹ GTGGAT ¹ TCGAT ¹ CTG ¹ CTAA ¹ TGTATAGAT ¹ TTGAAGAAGACGTGT ¹ TT	2104
Qy	836	ttaaacacatcacagaattcacaasaatcccaactatgttagtggccctccagttatggtg	895
Db	2105	TTACGAT ¹ CCCT ¹ CAGGAT ¹ TACA ¹ AA ¹ TTCA ¹ AA ¹ CTCA ¹ AA ¹ CTGCTAG ¹ TACCA ¹ CCCTAT ¹ TTTCA	2164
Qy	896	tttttggctaaagpccattagtcgatcaatacatgatttatcgactttaaocggaagttgct	955
Db	2165	TTCTTCGCCAAAGAC ¹ CTGTGAT ¹ TGC ¹ AAAT ¹ ACGAT ¹ TTATCTAAT ¹ TTACACGAAAT ¹ TGCT	2224
Qy	956	actggaggagctccttttagga ¹ aaagatctcgagaagcagtagcaagaaggttga ¹ aatta	1015
Db	2225	TC ¹ TGGGGGGCACCT ¹ TTTCGAAGAAG ¹ GTGGGGAAAGCGGTGCGAAACGC ¹ TTCCATCT ¹ T	2284
Qy	1016	cttggaaatcatacaaggatatgatttaactga ¹ aaacttgcctgcgtgtaatgattaccocct	1075
Db	2285	CCAGGGATACGACAAAGGATATGGCTCACTGAGACTACATCAGCTAT ¹ CTGATTACAC ¹ CC	2344
Qy	1076	cataatgctgtga ¹ aaacaggttcaactgggaagcccttgcatacatataaagctbaa ¹ g ¹ tt	1135
Db	2345	GAGGGGGATGATAACCGGGCGCGTCTGGTAAAGTTGTTCAT ¹ TTTGTGAAGCGAAG ¹ GT	2404
Qy	1136	ttagataacgctactgggaagcgttagaccaggagaagaagcga ¹ aaatagtctttcaa	1195
Db	2405	GTGGATCTGATACCGGGAAACGCTGGCGGTAA ¹ TACAGAGGCGGAAT ¹ TATGTGTCAGA	2464
Qy	1196	agtga ¹ aatgattatga ¹ agga ¹ tattacaacaatccgggaagcaactattgatactattgcac	1255
Db	2465	GGACCTATGATTATGTCGGT ¹ TATGAACAATCCGAAGCGACCAACGCTTGA ¹ TTGAC	2524
Qy	1256	aaagatggttggcttactctcggagatattggatatctacgcagaagatgga ¹ aatctcttt	1315
Db	2525	AAGGAT ¹ GGATGGGTACGAT ¹ CTGGAGACATAGCTTACTGGGACGAAGACGAAC ¹ CTCTTC	2584
Qy	1316	atggttgatcgattga ¹ agaactttataa ¹ tacaaggatcatcaggttgcgcctgtgaa	1375
Db	2585	ATAGTTGACCGCTTGAAGTCTTTTAATTAATACAANGATATCAGGTGGCCCCCGCTGAA	2644
Qy	1376	ctggaaaaatctgcttttacaacatccaagtattgctgtagcgggtgttactggagttccg	1435
Db	2645	TTGGAATCGATATTGTTTACAACACCCCAACATCTTTCGACGCGGGCGTGGCAGGTCT ¹ TCCC	2704
Qy	1436	gagcaatttgggtgacaatatacctgctgctgtgtgtgtgttagaatctgcaagacgctg	1495
Db	2705	GACGATGACCCCGGTGAAC ¹ TTCCCGCCGCGTGTGT ¹ TTTGGAGACACGGAAGACCGATG	2764
Qy	1496	actgaaaaaggaagttcaagattttattgcagocaaagtcactcca ¹ aaagaacatctcga	1555
Db	2765	ACGGA ¹ AAAGAGATCTGTGATTACGTGCGCAATCAAGTAAACAACCGGCA ¹ AAAGTTGGC	2824
Qy	1556	ggcgggtgcgtatttttagacagtatcccgaaagccctactgtgaa ¹ aactcatcaga ¹ aag	1615
Db	2825	GGAGGAGTGTGTCTGGACGACGATACCGAAGGTC ¹ TACCGGAA ¹ AACTCGACGCGAAGA	2884
Qy	1616	gagctccgaga ¹ aat 1629	
Db	2885	AAAAATCAGACAGAT 2898	

RESULT 6

RESULT b
US-09-334-765A-21

US-09-334-705A-21
; Sequence 21, Application US/09334765A

; Patent No. 6238893

; PATENT NO. 02360893
; GENERAL INFORMATION:

APPLICANT: Fallaux, Frits J.

APPLICANT: FILLIUX, FILLS J.
; APPLICANT: Hoeben, Robert C.

```

1  APPLICANT:  Bout, Abraham
2  APPLICANT:  Valeirio, Domenico
3  APPLICANT:  Van der Eb, Alex J.
4  TITLE OF INVENTION:  PACKAGING SYSTEMS FOR
5  TITLE OF INVENTION:  IN GENE THERAPY
6  FILE REFERENCE:  3833.2US
7  CURRENT APPLICATION NUMBER:  US/09/334,765A
8  CURRENT FILING DATE:  1999-06-16
9  PRIOR APPLICATION NUMBER:  US 08/793,170
10 PRIOR FILING DATE:  1997-03-25
11 PRIOR APPLICATION NUMBER:  PCT/NL96/00244
12 PRIOR FILING DATE:  1996-06-14
13 PRIOR APPLICATION NUMBER:  EP 95201728.3
14 PRIOR FILING DATE:  1995-08-26
15 PRIOR APPLICATION NUMBER:  EP 95201611.1
16 PRIOR FILING DATE:  1995-08-15
17 NUMBER OF SEQ ID NOS:  22
18 SOFTWARE:  Corel WordPerfect 8.0
19 SEQ ID NO 21
20 LENGTH:  5620

```

Query Match	31.48:	Score	554.4:	DB 4:	Length	5620:
Best Local Similarity	59.2%:	Pred. No.	2.7e-107:			
Matches 967:	Conservative	0:	Mismatches	661:	Indels	6:
					Gaps	1:

Qy	2	cag/cgaagactcttagggatc	aaatggaagaagaaacattataggca	tggaagagcgtcctt	61
Db	1265	cattccggtactgttggt	aaatggaagcgccaaacataaagaag	gccgcgcgcca	1324
Qy	62	cgfgatatgtccatccctg	ctcggcagagacaacattataccaat	catgtataaaattt	121
Db	1325	ttctatccctctagaggt	ggaacccgtggagcgcaactgcgataaggctatgaagaatac		1384
Qy	122	gcatacttttccgaagca	-----taatcgatgctctacaaatgaagtaatatcatat		175
Db	1385	gccctggttccctgaacaa	tgtctttacagatgcacatctcgaggaggaaacatcacgtac		1444
Qy	176	gctcaaatatttgaacaac	agtcgcgccttagctgttagtatagaaacaatat	tggtcttgaat	235
Db	1445	gcggataactctgaaat	gtccgttcggttggcagaagctatgaaacgatat	tggtcttgaat	1504
Qy	236	gaaacaaatgtgtgggt	gtatgcagtgaaacaaatacaaaactttttatccctgtcctt		295
Db	1505	acaaatcacagnatcgt	ctgtatgcagtgaaactctctcaattctttatgccggtgttg		1564
Qy	296	gctgcgttttactcttag	gaataccagtagcaacatcaaatgata	tgtaacacagatggagag	355
Db	1565	ggcgcgttattatcgg	agtgtscagttggccgcgcgaacacattataatgaaacgfgaa		1624
Qy	356	ttaactggttcatttga	atatatacaaaaaccaactatca	tgtttagtttcaagaagaagcactc	415
Db	1625	tgtctcaacagtatga	caatttcgcagctcacccgtagttgtgtgttccaaaaggggttg		1684

```
QY 416 ccgcttattctgagagtagacgacaaatctcaagttcttcattaaaaagtcgtagtatcgat 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1685 caaaaaattttgacgctgcacaaaaaattaccataatccagaaaaattattatcgat 1744
QY 476 agcatgtacacattaaaggcggtggaatgcgtatcatcacttttggcagcttatactgac 535
    || || || || || || || || || || || || || || || || || || || || ||
Db 1745 tctaaaaaggattaccagggtttcagtcgatgtacagcttcgtcacatctcatctacct 1804
QY 536 cacacctttgatccattgtcatttacacaaaaagatttttgatcccttgaaaaaatcgca 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1805 cccggttttaatagaatacagattttgacacagctccttgatcgtgacaaaaaatcgca 1864
QY 596 ttaattatgtcatctctggaaactaggattgctgaaggtgtagttactgagccataga 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1865 ctgataaagaattcctcgtgattactggtttacctaagggtgtggccctccgcataga 1924
QY 656 agctaaactataaagattcgtttcaatggcggtgatacccaatttatggcaactgcgtacggttcca 715
    || || || || || || || || || || || || || || || || || || || || ||
Db 1925 actgctgctgcagattctcgatgccagagatcctatttttggcaatcaaatcattccg 1984
QY 716 caaacatcaattcttctttagtagccgttccatcagctcctttggaatttttactacatta 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1985 gatactcgatttcaagtgctgtccatcccatcccggttttggaaatttttactacactc 2044
QY 776 tottactttgtagtaggacttaaggttgaattgttgaagaaatttgaggcgcacttttc 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 ggaatttgatagtgttttcgactgcttcttaattgtatagattttgaagaagcgtgttt 2104
QY 836 ttaaaaacatatagaaattacaaaaatccccactattgttagtggccccctccagttatggtg 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2105 ttacgatcccttcaggattacaaaattcaaatgcgttgcgtagtagtaccacacctatttca 2164
QY 896 ttttggctaaagccattgctgcatcaatacagatttatcagcttaacgagaagtgcct 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2165 tcttcgcacaaagcactgtgattgacaaatcacgatttatcctaattcacgaaattgct 2224
QY 956 actgaggagctcttttaggaaaaagatgctgcagagcagtagcagaaggttgaattta 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2225 tctggggcgccactcttctgaaaagaagtcggggagcggttgcaaaacgcttccatctt 2284
QY 1016 cctggaatacataaggatatggatttaactgaaaacttctgcgctgtgaattgataccacct 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2285 ccagggtacacaaaggatatgggtcactgagactacatcagctactctgtgatacacccc 2344
QY 1076 cataatgctgtgaacacaggttcaactggaagacccttgccatcatataaagctaaagt 1135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2345 gagggggatataaacccggcggtcgtgtaagttgttccatttttgaagcgaaggtt 2404
QY 1136 ttagataacgctacttggaagcgctagaccagagagaagaggcggaataatgctttcaa 1195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2405 gtggatctggataccgggaaaaacgctggcggttaatcagagagggcgaatttatgtctaga 2464
QY 1196 agtgaatgattatgaaggattattacaacatccggaagcaactattgatactattgac 1255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2465 ggacctatgattatgcggttatgtataacatcccggaagcgaacacgcttgattgac 2524
QY 1256 aaagatggttggtcttcattcttgagatatgtgatatcagacgaagatcgaaaattctttt 1315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 aaggatgagtggctacattctggagaacatagcttactgggacgaagacacactctctc 2584
QY 1316 atagttgctgattgaaagaacttattaaatacaagggatcatcaggttgcgcctgtgaa 1375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2585 atagttgaccgcttgaagtctttaaattaaatacaaaaggatatcaggttggcccccgctgaa 2644
QY 1376 ctggaaaaatcgtcttttaacacatccaagatttgctgtagtgggggtgttactggagttccg 1435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2645 ttggaaatcgattgtttcaacaccccccaacatcttcgacgcggcggtgtgacaggtctccc 2704
QY 1436 gacgaattgttggacaattacactgctgctgtgtgtgtgttagaatactggcgaagcgtgtg 1495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2705 gacgatgacgcgggtgaacttccgcgcgctgtgtgtgtgttttggagacacggaaaacgatg 2764
QY 1496 actgaaaaaggaaagttcaagatttttattgcagacaaagtcactccaacaaaagcatcttoga 1555
```

```
Db 2765 acggaataaagagatcgtgtgattacgtcccgagtcceagtaacaacccgcgaaaaagtgcgc 2824
QY 1556 ggcgggtcgtatttgtagacagtagtattccgaagaagccctactactggaaaaactcatcagaaaag 1615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2825 ggaggagtgtgttgtggacgaagtagccgaaggtcttaccggaaaaactcgacgcaga 2884
QY 1616 gagctcccgagaat 1629
    | || ||| |||
Db 2885 aaatcagagagat 2898
    || || ||| |||

RESULT 7
US-09-356-575E-21
; Sequence 21: Application US/09356575E
; Patent No. 6265212
; GENERAL INFORMATION:
; APPLICANT: Fallaux, Frits
; APPLICANT: Hoeben, Robert
; APPLICANT: Bout, Abraham
; APPLICANT: Valerio, Domenico
; APPLICANT: van der Eb, Alex
; APPLICANT: Schouten, Govert
; TITLE OF INVENTION: PACKAGING SYSTEMS
; FILE REFERENCE: 2578-3935US
; CURRENT APPLICATION NUMBER: US/09/356.575E
; CURRENT FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/793,170
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: PCT/NL96/00244
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: EP 95201611.1
; PRIOR FILING DATE: 1995-06-15
; PRIOR APPLICATION NUMBER: EP 95201728.3
; PRIOR FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 5620
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Derived from Adenovirus
US-09-356-575E-21

Query Match 31.4%; Score 554.4; DB 4; Length 5620;
Best Local Similarity 59.2%; Pred. No. 2.7e-107;
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;

QY 2 cagtgcagagacttttagggatcaaaatggaaagaagaacacattaggtcgtgagagcgctcct 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 cattccggtactgttggtaaaatggaaagaccccaaaaacataaagaagagcccgccgcga 1324
QY 62 cgtgatatagtcacactcctggctcgcgaggaacaaacattatcccaatcattgtataaaattt 121
    || || || || || || || || || || || || || || || || || || || || ||
Db 1325 tctatcctctagaggatggaaacgctggagacactgcataaaggtctatgaagaatac 1384
QY 122 gcactctttctcgaagcaaa-----taatcgatgctctatacaaaatgaagtaaatatcatat 175
    || || || || || || || || || || || || || || || || || || || || ||
Db 1385 gccctggttctctgaaacaaattgctttttacagatgcacatatcgcaggtgaaacatcacgac 1444
QY 176 gctcaaatatttgaaccagctggcgcttagctgttagtatagaaacataatggtccttgaat 235
    || || || || || || || || || || || || || || || || || || || || ||
Db 1445 gcggaactctgaaaatgtccggttcggttggcagaagctatgaaacgatatgggctgaat 1504
QY 236 gaaacaaattgttgggtgtatgcagtgaaacaaataaaaactttttttaaactcctgtcctt 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 acaaatcacagaatcgtctgtagtcagtgaaactctcttcaattctttatgccggtgtg 1564
QY 296 gctgctttatacttaggaataccagtagcaacatacaaatgatatactacagatggagag 355
    || || || || || || || || || || || || || || || || || || || || ||
Db 1565 ggcggttattatcggaggttgcagttgcgcccggaacgacattttataaagaacgtgaa 1624
```


PUBLICATION INFORMATION:

AUTHORS: de Wet, J. R.
AUTHORS: Wood, K. V.
AUTHORS: Deluca, M.
AUTHORS: Helinski, D. R.
AUTHORS: Subramani, S.
JOURNAL: Mol. Cell. Biol.
VOLUME: 7
PAGES: 725-737
DATE: 1987
us-08-857-352-22

Query Match 31.38; Score 552.8; DB 3; Length 1811;

Best Local Similarity 59.1%; Pred. No. 4.8e-107;

Matches 966; Conservative 0; Mismatches 662; Indels 6; Gaps 1;

Qy 2 cagtcgaagctttaggagatcaaaatgaagaagaaacattagcagtgagagcgctct 61
Db 73 CATTCGGGTACTGTGGTAAATGGAAGACGCCAAACATAAAGAAAGCGCCGCCA 132
Qy 62 cgtgatagtccttgcctgctgcgcggagacaacaattatatacaatcattgtataaattt 121
Db 133 TTCATCCTCTAGAGGATGGAACCGCTGGAGACCACTGCATAGGCTATGAGAGATAC 192
Qy 122 gcatcttttccctgaagca-----taatcgatgctctacaaatgaagtaatatcat 175
Db 193 GCCCTGGTCTCGTAACAATTTGCTTTACAGATGCACATATCGAGGTGAACATCACGTPAC 252
Qy 176 gctcaaatatttgaacacagctgcgccttagctgttagtagtagaacaatatggcttgaat 235
Db 253 CGGNAATCTCGAANAATGCTCGGTGGCAGAGCTATGAACGATATGGGCTGAAT 312
Qy 236 gaacaaatgttgggtgtatgcagtgaaacaaataaaacttttttaactcctgtcctt 295
Db 313 ACAATACAGAAATCGTGTATGAGTGAACAACTCTCTCAATCTCTTATGCGGTGTG 372
Qy 296 gctgctttatacttaggaataccagtagcaacatacaaatgatagtagtacacagatggagag 355
Db 373 GCGCGTATTTATTCGGAGTTGCGAGTTGCGCGCGGAGACGACATTTAATGAACGTGAA 432
Qy 356 ttaactggtcattgaaatatacaaaacaaactatcatagtttagttcaagaagaacacac 415
Db 433 TTGCTCAACAGATGAACATTCGACGCTACCGGTAGTGTGTTGTTCCAAAAGGGGTG 492
Qy 416 cgcgtttatctgagtagtacgaacaaataaagtttcaataaaagtcgttagtctatcgat 475
Db 493 CAAAAAATTTGAACGTGCAAAAAAATTTACCAATAATCCAGAAAAATTTATTATCATGGAT 552
Qy 476 agcatgtacgacattaaatggcgtggaatgcgtatctactcttctgttgacggttatactgac 535
Db 553 TCTAAACGGATTAACAGGGATTTTCAGTCGATGTACAGGTTCGTACATCTCATCTACCT 612
Qy 536 cacaccttggatccatgtctatttaccacaaagatttggatcccttgtaaaaaatcgca 595
Db 613 CCGGTTTTAATGAATACGATTTGTACACAGTCCCTTTGATCGTGCACAAAACAAATTCGA 672
Qy 596 ttaattatgcatcatctggaacaaactggattgctgaagggtgtagtagtagcagccataga 655
Db 673 CTGATAATGAATTCCTCTGGATCTACTGGTGTACCTAAGGGTGTGGCCCTTCGCCATAGA 732
Qy 656 agtctaactataaagatttgcttagcagagggatccatttatggcactcgtacgggtcca 715
Db 733 ACTGCTCGGTGAGATTCGCATGCCAGATCCTATTTTGGCAATCAATCATTCCTCG 792
Qy 716 caaacatcaattcttctcttagtaccgtttccatcatcgtcttggatgttttactactatta 775
Db 793 GATATCGGATTTTAAGTGTGTTCCATCCATCCATCAGGTTTGGATGTTTACTACATC 852
Qy 776 tcttactttgttaggacttaagttgttaattgttgaagaatttgaaggcgacattttc 835
Db 853 GGATATTGATGTGGATTTTCGATCGTCTTTAATGTATAGATTGAAGAGAGCTGTTT 912

Qy 836 taaaaaccatacagaattacaaaatccccactattgtagtgccctccagttatggtg 895
Db 913 TTACGATCCCTTCAGGATTCAAAATTCAAAGTGGTGTAGTACCAACCCCTATTTTCA 972
Qy 896 tttttggctaaagccctatttagtcgatacaatcacgatttatcgagcttaacggaagttgct 955
Db 973 TTCTTCGCCAAAAGACACTCTCATTTGACAAATACGATTTATCTAATTTACAGAAATGCT 1032
Qy 956 actggaggagctccttttaggaaaagatgctgcagaagcagtagcaaaagaggttgaaaa 1015
Db 1033 TCTGGGGCGCACCTCTTTTCGAAAAGAGTCGGGAAGCGTTGCAAAACGCTTCATCTT 1092
Qy 1016 cctggaatcatacagaagtagtattaaactgaaacttgcgtgctgtaataatcaccct 1075
Db 1093 CCAGGATACCAACAGGATATGGGCTCACTGAGACATCATCAGCTATTTCTGATTACACC 1152
Qy 1076 cataatgctgtaaacacaggttcaactggaagaccccttgcatacatataaagctaaaagt 1135
Db 1153 GAGGGGATGATTAACCGCGCGCTCGTAAAGTTGTTCCATTTTGTGAAGCGAAGGTT 1212
Qy 1136 ttgataacgctactggaagcgcttaggaccaggaagaaagagcgaaatatgcttttcaa 1195
Db 1213 GTGATCTGGATACCGGAAAACGCTGGCGTTAATCAGAGAGCGGAATTTATGTGTGACA 1272
Qy 1196 agtgaatatgattgaagaagattatacaacaatccggaagcaactattgatactattgac 1255
Db 1273 GGACCTATGATTTATGTCGGGTATGTAAACAATCCGAAGCGCAACGCTTGAATTGAC 1332
Qy 1256 aaagatggttggcttcatctcggagatatggaattatcgacgaagatggaaatttctt 1315
Db 1333 AAGATGATGGCTACATTTCTGGAGACATAGCTTACTGGGACGAAGACAGACATCTTTC 1392
Qy 1316 atagttgatcagattgaagaacttattataaatacagaaggaatatcaggttgcgcttctgaa 1375
Db 1393 ATAGTTGACCGCTTGAAGCTTTAATTAATACAAAGGATATCAGTGGTGGCCCGCTGAA 1452
Qy 1376 ctgaaaaatctgcttttacaacatccaaagtattgctgacgggtgttactggaagttcog 1435
Db 1453 TTGGAATCGATATTGTTACAACACCCCAACATCTTCGACGCGGGCTGGAGGCTTCTCC 1512
Qy 1436 gacgaatttggtagaattacactgctgctgtgtgtgtgttagaactctggcaagacgctg 1495
Db 1513 GACGATGACGCGGTGAACCTCCCGCCCGCTGTGTGTTTGGAGCACGGAAGACGATG 1572
Qy 1496 actgaaaagggaagtccaagattttatgtcagcacaaagtcactcccaaaagacatcttca 1555
Db 1573 ACGGAAAAGAGATCGTGGATTTACGTCGCCAGTCAAGTAACAACCGCGAAAAGTTGGCG 1632
Qy 1556 ggcggtgtcgtattttagacagatttccgaagccctacttgaaactcatcagaagaag 1615
Db 1633 GGAGGAGTTGTGTTTGGACGAGTACCGAAAGGCTTTACCGGAAAACACTCGACGCAAGA 1692
Qy 1616 gagctcccgagaaat 1629
Db 1693 AAATCAGAGAT 1705

RESULT 10

US-08-862-431-32

; Sequence 32, Application US/08862431

; Patent No. 6120994

; GENERAL INFORMATION:

; APPLICANT: TAM, SHUI-PANG

; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P. L. C.

; STREET: 1100 NEW YORK AVENUE, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: US

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/862,431
FILING DATE: 23-MAY-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kim, Judith U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 1669.0020000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 5789 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
US-08-862-431-32

Query Match 31.3%, Score 552.8; DB 3; Length 5789;
Best Local Similarity 59.1%; Pred. No. 5.8e-107;
Matches 966; Conservative 0; Mismatches 662; Indels 6; Gaps 1;

```
Qy 2 cagtgcaagacttaggacatacaaaatgaagaagaacacattagggatggagagcgctct 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 CATTCGGGTACTGTGGTAAATGAAGACGCCCAAAACATAAAGAAAGGCCCGCGCCA 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 62 cgtgatagtccactcctggctcggcagagcaacaaattataccaatcattgtataaatt 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 TTCATCTCTAGAGGATGGAACCGCTGGAGACCAACTGCATAAGGCTATGAAGAGATAC 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 122 gcacttttctctgaagcaa-----taacgatgctcacaacaaatgaagtataatcat 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GCGCTGGTTCCTGGAAACAATTGCTTTACAGATGCACATATCGAGGTGAACATCACGTAC 426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 176 gctcaaatattgaaacagctgcgcttagctgttagtatagaacaatatggcttgaat 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 GCGGAATACTTCGAAATCTCGCTTGGTGGCAGAAGCTATGAAACAGATATGGGCTGAAT 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 236 gaaaacaattgttggtgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ACAATACACAGATCGTGTATGACGTGAAACACTCTCTCAATCTTTATGCGCGGTGTG 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 296 gctgcttatacttaggaataaccagtagcaacaatacaaatgatatgacacagatggag 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 GCGCGGTATTATTCGGAGTTGCAAGTTCGCGCGCGGAACGACATTTAATGAACGTGAA 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 356 ttaactggtcatttgaatatcaaaaacaaactalcalgttttagttcaaaagaagcactc 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 TTGCTCAACAGATGAACATTTTCGAGCGCTACCGTAGTGTGTGTTTCCAAAAGGGGTG 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 416 ccgcttattctgagagtagcagaaatactaagtctcaataaaagtcgttagttatcgat 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 CAAAAAATTTGAAGCTGCAAAAAAATTAACCAATAATCCAGAAAAATTTATCATGTGAT 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 476 agcatgacacattaaatggcgttgaatgctgatatctaactttgttgcagttatactgac 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 727 TCTAAACGGATTACAGGGATTTTCAGTCGATGTACAGCTTCGTACATCTCATCTACCT 786
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 536 cacacatttgatccattgtcattacacaaaagaatttttgatcccttgaaaaaatcgca 595
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 CCCGGTTTAAATGAATACGATTTTGTATCCAGAGTCCTTTGATCGTGCACAAACAAATTGCA 846
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 596 ttaattatgcatcatcctggaaacactgattgcttaagggtgtagtgagccataga 655
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 847 CTGATAATGAATTCCTCTGGATCTACTGCGTTACCTAAGGGTGTGGCCCTTCCCGATAGA 906
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 656 agtctaactaagaattcgttcatagcaggggaggtccatttatggcaotcgtacggttcca 715
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11
US-08-862-431-31
; Sequence 31, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:

```
Db 907 ACTGCTCGTCAGATTCGTCATGCCAGAGATCCTATTTTGGCAATCAAAATCATTCGG 966
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 716 caaacataaattcttctcttagtcagtcgcttccatcatgcttctgaatttactacatta 775
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 967 GATACGCGATTTTAAAGTGTGTTCATTCCTACCGGTTTGGAAATTTTACTACATC 1026
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 776 tcttactttgtagtaggacttaaggattgaattgtgaagaattttgaggggcgacatttc 835
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1027 GGATATTTGATATGTGGATTTTCGAGTCCTCTTAATGTATAGATTTGAAGAAGAGCTGTT 1086
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 836 ttaaaacacatacagaattacaaaatccccactattgttagtggccccctccagttatg 895
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1087 TTACGATCCCTTCAGGATTACAAAATTCAAAAGTGGTGTCTAGTACCAACCGCTATTTCA 1146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 896 tttttggctaaaagcccattagtgcatacagatttatcgagcttaacaggaagtgc 955
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1147 TTCTTCGCCAAAAGCACCTCTGATTGACAAATACGATTTTATCTAATTTACAGAAATGCT 1206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 956 actggagagctccttttaggaaaagatgtcgcagagcagtagcagaaggtttgaaatta 1015
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1207 TCTGGGCGCACCTCTTTTCGAAAGAACTCGGGGAAGCGTTGCAAAACGCTTCCATCTT 1266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1016 cctggaatcatacagaagatatggattaaactgaaacttgcgcgtgtaattgattaccct 1075
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1267 CCAGGGATACGACAAAGGATATGGGCTCACTGAGACTACATCAGCTATTTCTGATTACACCC 1326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1076 cataatgcttgaaacaggttcaactggaagacccttgccatcattaaagcttaagtt 1135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1327 GAGGGGATGATAAACCGGCGCGCTCGTAAAGTTGTTCATTTTTTGAAGCGAAGTT 1386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1136 ttagataacgctactctgggaagcgctaggaccagaggaagggcgaaatgctcttcaa 1195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1387 GTGGATCTGGATACCGGGGAACCGCTGGCGTAAATCAGAGAGGGAATTTATGTCTAGA 1446
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1196 agtgaatgattatgaagagatatatacaaatccggaaagcaactattgatactattgac 1255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1447 GGACCTATGATTATGCTCCGGTTATGTAACAATTCGGGAAGGACCAACGCGCTTGATTGAC 1506
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1256 aaagatggttgcttcattctggagatatattgatatattacgaagatggaaatttcttt 1315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1507 AAGGATGGATGGCTACATTTCTGGAGACATAGCTTAACTGGGAGCAAGCAACACTTCTTC 1566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1316 atagtgtatcgtattgaagaacttattataacaaaggtatatacaggttgcgctgtcga 1375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1567 ATAGTTGACCGCTTGAAGCTTTAATTAATACAAAGGATATCAGGTGCCCCCGCTCAA 1626
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1376 ctgaaaaatctgcttttacaacatccaagtattctgctgtagtgcgggtgctactggag 1435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1627 TTGGAATCGATATTGTTACACACCCCAACATCTTCGACGGGGGCTGGCAGGTCTCCC 1686
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1436 gacgaatttgtagcaaatattacctgctgtgtgtgtgttagaatctggcaagacgctg 1495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1687 GACGATACGCGGGTGAACCTTCCCGCCCGCTGTGTGTGTTGGACGCGAAACGATG 1746
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1496 actgaaaaggagttccaagattttattgcagcacaagtcaactcccaacaaagcattcga 1555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1747 ACGGAAAAGAGATCGTGGATTACCTGCCAGTCAAGTAAACACCGCGCAAAAAGTTGCGC 1806
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1556 ggcgggtgctgtattgttagcagattatccgaaaggccctactggaaaaactcatcagaag 1615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1807 GGAGGAGTGTGTTGTGGAGCAAGTACCGAAMAGTCTTACCGGAAAAACTCGACGCAAGA 1866
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1616 gagctccgagaaat 1629
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1867 AAAATCAGAGAGAT 1880
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Qy 1436 gacgaatttgggtgacnataactctgccttgctgtgttgttagaatctggcagaagcgtg 1495
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1691 GAGCATGACGCCCGGTGAACCTCCCGCCGGCTTGTGTGTGTGGAGCACGGAAAGACGATG 1750

Qy 1496 actgaaaaagaagtctcaagattttattcgacacaagtcacctccaacaaagcatcttcga 1555
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1751 ACCGAAAAGAGATCCTGGATTACGTGCCAGTCAAAGTAACAACCGGAAAAAAGTTGC GC 1810
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1556 ggcggtgtcgtattttagacagatctccgaaaggccctactcggaaaaactcatcacgaag 1615
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1811 GGAGGAGTGTGTGTTGTCACCAACTACCGAAAGTGCTTACCAGGAAAAAATCGACGCAAGA 1870
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1616 gagctccgagaaat 1629
|| | ||||| ||

Db 1871 AAAATCAGAGAGAT 1884
|| | ||||| ||

RESULT 13

US-08-862-431-30

; Sequence 30, Application US/08862431

; Patent No. 6120994

; GENERAL INFORMATION:

; APPLICANT: TAM, SHUT-PANG

; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

; STREET: 1100 NEW YORK AVENUE, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: US

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/862,431

; FILING DATE: 23-MAY-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kim, Judith U.

; REGISTRATION NUMBER: 40,679

; REFERENCE/DOCKET NUMBER: 1669.0020000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5793 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

US-08-862-431-30

QY	176	gctcaaatatttgaaccagctgcgcgtcttagctgttagtatagaacaatcatgcttgaaat	235
DB	431	CGCGAATACTTCGAAATGTCGCTTCGGTTGGCAGAAGCTATCAAAACGATATCGGCTGAAT	490
QY	236	gaaaaaactgtgtgggtgtatgcagtgaaaaacaataaaaactttttaactcgtctcctt	295
DB	491	ACAAATCACAGNAATCGTGTATGTCAGTGAATACTCTCTCAATCTTATATGCGGTGTG	550
QY	296	gctgctttactacttaggaataccagtagcaacaatcaaatgataatgtacacagatggagag	355
DB	551	GGCGGCTTATTTATCGGAGTTCCAGTTGGCCCGCGCAAGCACATTTATTAATCAACGTGAA	610
QY	356	ttaactggctatttgaatatatacaaaaacaacta toatgttttagtttcaagaagaacactc	415
DB	611	TTGCTCAACAGTATGAACATTTCCGAGGCTACCGTAGTGTTTTGTTCCTCCAAAAGGGTTG	670
QY	416	cgcgttatctcagagtagcacgaaaaactcaagtcttcatataaaagctgtagtattcgat	475
DB	671	CAAAAAATTTTGAACGTGCAAAAAAATTTACCAATATCCAGAAATTTATTTATCATGGAT	730
QY	476	agcatgtacgacattaaatggcgttgaatgcgtacttaccttttgttgcagcttatactgac	535
DB	731	TCTAAACGGATTACCAGGGATTTTCAGTCGATGTACACGTTTCGTACATCTCATCTTACCT	790
QY	536	cacacctttgatccattgtcatttacacaaaagatttttgatcccoottgaaaaaatcgcac	595
DB	791	CCCGGTTTAAATGAATACGATTTTGTACAGAGTCTCTTTGATCGTACACAAATTTGCA	850
QY	596	ttaattatgtcatcctcgtgaacaactggatgtgcttaagggtgttagtactgagccataga	655
DB	851	CTGATTAATGAATTCCTCTGGATCTACTTGGGTTACCTTAAGGGTGTGGCCCTTCGCCATAGA	910
QY	656	agctcaactataagattcgttcattagcagggatcccatatta tggcactcgtcaggttcca	715
DB	911	ACTGCCTCGTCAGATTCGCGATGCCAGAGATCCTATTTTGGCAATCAATCATCTCCG	970
QY	716	caaacatacaattcttccttagtaacogttcccatcaltgcctttggaaatgtttactacatca	775
DB	971	GATATGGGATTTTAAAGTGTGTTCCATTTCCATTCACGGTTTGGATGTTTACTACATC	1030
QY	776	tcttacttttagtaggacttaagttgttaatgttgaagaaattttagggcgcaacttttc	835
DB	1031	GGATATTGTATGTGGATTTCGAGTCGTCTTAATGTATAGATTTTGAAGAAGAGCTGT	1090
QY	836	ttaaaaaccatacagaattcacaaaatccccactattttagtggccctccagttatcggtg	895
DB	1091	TTACGATCCCTTCAGGATTACAAATTCAAAGTCGCTGTCTAGTACCAACCTTATTTTCA	1150
QY	896	tttttggctaaaagccccattagtcgatacaatcagattttatcagcgttaacggaagtgt	955
DB	1151	TTCTTTCGCAAAAGCACCTGTGATGTACAAATACGATTATCTAAATTTACACGAAATTTGT	1210
QY	956	actgaggagctcctttaggaaaaagatgtcgcagaagcagtagcaaaaggttgaataat	1015
DB	1211	TCGTGGGGCGCACCTTTTCGAAAAGAAGTCGGGAAGCGGTTGCAAAACGCTTCCATCTT	1270
QY	1016	ccttgaatcatacaaggaatagtttaactgaaactgctgcgctgtaatgattacccct	1075
DB	1271	CCAGGGATACGACAGGATATGGGCTCACTGACATACATCAGCTATTCTGATTACACCC	1330
QY	1076	cataatgctgtgaaaacaggttcaactggaaagacccttgcccatacattaagagctaaagtl	1135
DB	1331	GAGGGGATGATAAACCGGGCGGTTCGCTAAAGTTGTTCCTATTTTGTGAAGCGAAGGTT	1390
QY	1136	ttagataacgctacttgggaagcgcttaggacccagagaaaaggcgaaaaatagctttcaa	1195
DB	1391	GTGGATCTGGATACCGGGAAAACGCTGGCGGTTAATTCAGAGAGCGCAATATGTGTGACA	1450
QY	1196	agtgaatgattataaaggatattacaacaatcccggaagcaactattgatactatgtac	1255
DB	1451	GGACCTTATGATTATGTCCGGTTATGTAAACAATTCGGAAGCGACCAACGCCCTTGATTGAC	1510
QY	1256	aaagatggttggtcttcttccttgagatatattggaattatacgaagaatggaaattcttt	1315

Db	1536	AAGGATGGATGGCTACATTC	TCTGGAGACATAGCTTTACTGGGACGAGACGAAACACTTCTTC	1595
Qy	1316	atagttgatcgattgaagaactatt	aaatacaaggatatacaggttgcgcctgctgaa 	1375
Db	1596	ATAGTTGACCGCTTGAAGTCTT	TAAATTAATACAAGGATATCAGGTGGCCCGCGTGAA 	1655
Qy	1376	ctggaaaaatctgcttttacancat	ccaagtattgctgatcggggtgttactggagttccg 	1435
Db	1656	TTGGAATCGATATTTGTATCA	ACACCCCAACATCTTCGACGCGGGCGTGGCAGGTCTTCCC 	1715
Qy	1436	gacgaatttggtggacaatactac	ctgctgctgtgttggttagaatactggcagaacgctg 	1495
Db	1716	GACGTGACGCCGGTGAAGTTC	CCGCCGCGTGTGTTGTTTGGAGCAGCGAAAGACGATG 	1775
Qy	1496	actgaaaaggagattcaagatttt	atgacgacaaagtcactcaacaaagcatcttcga 	1555
Db	1776	ACGMAAAGAGATCTGTGATTC	GCGCATCAAGTTAACACCGCGAAMAGTTTCGC 	1835
Qy	1556	ggcgggtgcgtattttgacagat	ttccgaaggccctactggaaaactcatcagaag 	1615
Db	1836	GGAGGAGTTCTGTTCTGGAC	GAAGTACCGAAAGGCTTACCGGGAAACACGACGCAAGA 	1895
Qy	1616	gagctccgagaat	1629	
Db	1896	AAAATCAGAGAT	1909	

```

RESULT 15
US-08-536-559A-2
; Sequence 2, Application US/08536559A
; Patent No. 5994061
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Constructs and Methods for Screening for
; TITLE OF INVENTION: Increased Expression of Human apo AI Gene
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/536,559A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5819 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA (plasmid) pGL2(apo
; DESCRIPTION: AI-250)luc"
US-08-536-559A-2

```

	Query Match	31.3%	Score 552.8	DB 2	Length 5819
	Best Local Similarity	59.1%	Pred. No. 5.8e-107		
	Matches 966	Conservative	0	Mismatches 662	Indels 6
Qy	2	cagtcgaagcattttagggatccaataatgggaagaaacatttaggcattggagagcgtcct	61		
Db	277	CATTCCGGTACTCTTGTTAAATGGAAGACGCCAAAAACATAAGAAAGGCCGCGGCCA	336		
Qy	62	cgigtatagtcctatcctggctcggcagacacaaattatcaccaatcattgtataaaattt	121		
Db	337	TTCTATCCTCTAGAGNTGMAACCGCTGGAGAGCAACTGCATATAGGCTATGGAAGAGATAC	396		
Qy	122	gcattcttttctctaagaaca-----taatcgatgctgtacatacaaatgaagtaatacatat	175		
Db	397	GCCCTGGTTCTCTGGACAAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACCCTAC	456		
Qy	176	gctcaaatatttgaaccaggctgccgcttagctggttagtagaacacaatattggcttgaat	235		

Db	457	GGCGAATAC	TCTCGAAATG	TCCTCGGTTG	CGCAGAGCTAT	GAACGATAT	GGCTGAAT	516
Qy	236	gaacaacattg	tgtgtgtgtat	tcagtgaaacaa	lataaacctt	tttttaactgt	ctctt	295
Db	517	ACAAATCAC	AGAAATCGT	ATGCGATG	CAAGTGA	AAACTCTCT	CAATCTTTAT	576
Qy	296	gctgttatact	taggaatacc	agtagcaacat	caaatgatat	gtatgtacacagat	ggagag	355
Db	577	GGCGGTTAT	TTATATCGG	AGTTCAGT	TGCGCGG	CAACGATAT	TATGAACGTGA	636
Qy	356	ttaactggct	cattgaatat	tcaaaaccaac	tatcatgtt	tagttcaagaag	cacatc	415
Db	637	TTGTCTAAC	AGTATGAAC	ATTTCCGAG	CCCTACG	TAGTGTTTG	TTTCCAAAAGGG	696
Qy	416	cogctattct	gagagtagc	agaaactaa	gttcaatt	taaaaaagtc	agtagtattc	475
Db	697	CAAAAAAT	TTTCAACG	TCGCAAAAA	AAATTTAC	CAATATCC	AGAAAATTTAT	756
Qy	476	agcatgtac	agcaataa	tgogttgaat	gcgtatct	tacotttgt	gcaggtatata	535
Db	757	TCTAAAC	CGGATTTAG	CAGGGATTT	CAGTCGAT	GTACACGTT	CGTCAATCTCAT	816
Qy	536	cacacctt	gatccatt	gtcatttac	caacaaag	attttagtcc	cttcaaaaaat	595
Db	817	CCCGGTTT	TTAATGA	TACAGTTT	TGTAC	CAGAGTCCTTT	TGATCGT	876
Qy	596	tcaattat	gtcatcat	cttggaac	actggatgc	cttaagggt	gttagtact	655
Db	877	CTGATAAT	GAATTTCT	CTCGGATCT	ACTGGGTT	ACCTAAGG	GTGGCCCTTCCG	936
Qy	656	agcttaact	ataagatt	ctgtcat	agcaggat	ccccattat	tgccactgc	715
Db	937	ACTGCTG	CGTCAGAT	TTCTCGCAT	GCCAGAGAT	TCCTATTTT	TGGCAATCAAA	996
Qy	716	caaacatca	attcttct	cttagt	atccgttc	atcatgcct	tgtggaa	775
Db	997	GATAC	TGCGATTTT	TAAAGTGT	TGTCATTT	CCATACAG	GTTTTTGG	1056
Qy	776	tctactt	gttagt	agactta	aggttca	atgtga	agaaatttg	835
Db	1057	GGATATT	TCATATG	YGGAATTT	CGAGTC	GTCTTANT	GTATAGATTT	1116
Qy	836	ttaaaacc	atacaga	attacaaa	atcccc	actatgt	tagtggccct	895
Db	1117	TTACGAT	CCCTTCAG	GATTTAC	AAAAATTC	AAAGTGC	GTGTAGTAC	1176
Qy	896	titttggct	taaaagcc	attagtc	ataatc	agattat	catcagc	955
Db	1177	TTTCTG	CGCAAAAG	CACTCTG	ATTC	CAAAATAC	GATTTATCT	1236
Qy	956	actgagag	gtccctt	tagga	aaaagat	gtcgc	agaagcag	1015
Db	1237	TCG	GGGGCG	CACTCTTT	CGAAA	GAAGTC	GGGGAAC	1296
Qy	1016	cctggaat	catatac	aggtat	ggattaa	actgtc	gtcgctga	1075
Db	1297	CCAGGG	ATACG	ACAGGAT	ATGGGCT	CACTG	AGACTACAT	1356
Qy	1076	cataat	gtgtg	aaaaac	aggttca	actgtga	agaccott	1135
Db	1357	GAGGGG	ATGAT	AAACCG	GGCGGT	CGGTAA	GTGTTC	1416
Qy	1136	ttagata	acgtact	gtggag	gcgtc	tagcacc	agagaga	1195
Db	1417	GTGGAT	CTGGAT	ACCGGG	AAAAAC	CGCTGG	GGGTTAAT	1476
Qy	1196	agtgaa	atgatt	gaagga	tattaca	caatcc	gggaagc	1255
Db	1477	GGACCT	ATGATTA	GTCCCG	TTATGT	TAACAA	TCCGAG	1536
Qy	1256	aaagat	gtgttg	cttca	ttctgt	gagata	ttac	1315
Db	1537	AAGGAT	GATGGCT	ACATTTCT	GGAGAC	ATAGTCT	TACTGG	1596

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2002, 19:55:51 ; Search time 296.36 Seconds
(without alignments)
10225.246 Million cell updates/sec

Title: us-09-993-874-1
Perfect score: 1765
Sequence: 1 tcagtgaagacttttaggga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 347872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	575.2	32.6	2193	22 AAF24923	Complement of Fc e
2	575.2	32.6	2268	22 AAF24920	Nucleotide sequenc
c 3	575.2	32.6	2268	22 AAF24921	Complement of Fc e
4	573.8	32.5	1638	22 AAF24919	Nucleotide sequenc
5	571.6	32.4	1642	20 AAX32854	Mutant luciferase
6	571.6	32.4	1642	20 AAS00896	Beetle cDNA encodi
7	570	32.3	1639	20 AAX32847	Mutant luciferase
8	568.4	32.2	1639	20 AAX32844	Mutant luciferase
9	566.8	32.1	1639	20 AAX32843	Mutant luciferase

10	566.8	32.1	1639	20 AAX32845	Mutant luciferase
11	566.8	32.1	1639	22 AAS00889	Beetle cDNA encodi
12	565.2	32.0	1639	20 AAX32846	Mutant luciferase
13	565.2	32.0	1639	22 AAS00886	Beetle cDNA encodi
14	565.2	32.0	1639	22 AAS00887	Beetle cDNA encodi
15	563.6	31.9	1639	22 AAS00888	Beetle cDNA encodi
16	562.4	31.9	1639	20 AAX32849	Mutant luciferase
17	560.8	31.8	1639	20 AAX32852	Mutant luciferase
18	560.4	31.8	1639	22 AAS00885	Beetle cDNA encodi
19	560.4	31.8	1639	22 AAS00885	Beetle cDNA encodi
20	559.8	31.7	4987	16 AAO86799	Vaccinia virus vec
21	559.2	31.7	1639	20 AAX32853	Mutant luciferase
22	559.2	31.7	1639	22 AAS00891	Beetle cDNA encodi
23	558.8	31.7	1639	22 AAS00898	Beetle cDNA encodi
24	557.6	31.6	1639	20 AAX32850	Mutant luciferase
25	557.6	31.6	1639	20 AAX32851	Mutant luciferase
26	556	31.5	1639	22 AAS00894	Beetle cDNA encodi
27	555.6	31.5	1639	22 AAS00899	Beetle cDNA encodi
28	552.8	31.3	1798	20 AAX06815	Firefly luciferase
29	552.8	31.3	1811	15 AAO38732	Encodes firefly lu
30	552.8	31.3	5620	18 AAT48630	Ad5-ITR, CMV-lucif
31	552.8	31.3	5620	22 AAF30233	Minimal adenovirus
32	552.8	31.3	5789	20 AAX08779	pGL2-Promoter gene
33	552.8	31.3	5791	20 AAX08778	pGL2-(GST-ARE)SV40/
34	552.8	31.3	5793	20 AAX08776	pGL2(apo AI-ARE)SV
35	552.8	31.3	5793	20 AAX08777	pGL2(apo AI-mARE)S
36	552.8	31.3	5818	21 AAZ38352	Plasmid pGL2 (apoA
37	552.8	31.3	5819	20 AAX08775	pGL2(apo AI-250)lu
38	552.8	31.3	5819	20 AAX08774	pGL2(apo AI-250)lu
39	552.8	31.3	5819	21 AAZ38351	Plasmid pGL2 (apoA
40	552.8	31.3	5938	21 AAZ38353	Plasmid pGL2 (4xap
41	552.8	31.3	6092	20 AAX08773	pGL2(apo AI-491)lu
42	552.8	31.3	6092	21 AAZ38350	Plasmid pGL2 (apoA
43	552.8	31.3	6170	16 AAO98929	Vector p19LUC sens
44	552.8	31.3	6254	16 AAO98916	TGF-beta responsiv
45	552.8	31.3	6464	16 AAO98933	Vector HBVLuc. Sy

ALIGNMENTS

RESULT 1
AAF24923/C
ID AAF24923 standard; DNA; 2193 BP.
XX
XX AAF24923;
XX
XX 30-APR-2001 (first entry)
XX
XX Complement of Fc epsilon receptor-luciferase fusion protein DNA.
DE
XX
XX Fc epsilon receptor; Fc epsilon receptor; immunoglobulin E; IgE; atopic disease;
KW luminescence inducing protein; allergy; hyper IgE syndrome; chimera;
KW internal parasite infection; B cell neoplasia; luciferase; ss.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Photuris pennsylvanica.
XX
XX WO200104310-A1.
XX
XX 18-JAN-2001.
XX
XX 13-JUL-2000; 2000MO-US19070.
XX
XX 13-JUL-1999; 99US-0143612.
XX 02-MAR-2000; 2000US-0186412.
XX
XX (HESK-) HESKA CORP.
XX (PROM-) PROMEGA CORP.
XX
XX Weber ER, Wood KV, Hall MP;
XX WPI; 2001-103082/11.
XX

Qy	1400	ccaagtattgtgatcggggtgttacttgaggtctccgcgcgaatttgggtgacaattacct	1459
Db	1381	ccgtatttattgtatccggcgcttacttggtatcaccggatgaagcgcgggcgagcttcca	1440
Qy	1460	gctgcttgttgttgaatactggccaagcgcgtcgtcgtgaaaaggaagttcagatctt	1519
Db	1441	gctgcagggtgtgtcagtcacagactggaaaatactcaacgcacaaaatcgtaaaaaattt	1500
Qy	1520	attgcgacacaagtcactccaacaagaatcttcgagcggcgtcgtatattttagacagt	1579
Db	1501	gtttccagtcagaatttcaacagccaaatggctacgtggtgggtgaaaattttggatgaa	1560
Qy	1580	attccgaagaagccactactggaagaactcatcagaagaagcgtccgagaaaattttgccag	1639
Db	1561	attcccaaaaggtacaactggaataattcacagaaaagttcaagcacaatgtttgaaaaa	1620
Qy	1640	cgagcaccaaa	1650
Db	1621	cacaaatctaa	1631
RESULT 5			
AA32854			
ID	AA32854 standard; DNA; 1642 BP.		
XX	AA32854;		
DT	28-JUN-1999 (first entry)		
XX	Mutant luciferase encoding DNA.		
DE	Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;		
KW	luciferase; genetic reporter; enzyme immobilisation; mutant; ss.		
XX	Photuris pennsylvanica.		
OS	Synthetic.		
XX	W09914336-A2.		
XX	25-MAR-1999.		
XX	18-SEP-1998; 98WO-US19494.		
PR	19-SEP-1997; 97US-0059379.		
PA	(PROM-) PROMEGA CORP.		
PI	Hall MP, Wood KV;		
XX	WPI; 1999-229538/19.		
DR	P-PSDB; AAY06861.		
PT	New mutant luciferase enzymes with increased stability		
XX	Claim 16; Page 88; 167pp; English.		
CC	The invention relates to a mutant luciferase enzyme that have increased		
CC	thermostability. The mutant luciferase is obtained by mutating a		
CC	polynucleotide sequence encoding a beetle luciferase and selecting		
CC	expression products having increased thermostability. The new		
CC	luciferases can be used in ATP assays, as luminescent labels for nucle-		
CC	acids, proteins or other macromolecules, as genetic reporters, in enzyme		
CC	immobilisation, as hybrid proteins, in high temperature reactors, and		
CC	in luminescent solution. The improved thermostability allows storage of		
CC	luciferases without altering its activity, and improves reproducibility		
CC	and accuracy of assays using the new luciferases. The novel luciferases		
CC	can also have improved brightness of the luciferases. The novel luciferases		
CC	and luminescence signal. Sequences AA32843-856 represent DNA molecules		
CC	encoding the mutant luciferase enzymes.		
XX	Sequence 1642 BP; 559 A; 255 C; 322 G; 506 T; 0 other;		

Query Match 32.4% Score 571.6; DB 20; Length 1642;
Best Local Similarity 59.9%; Pred. No. 9.5e-102;
Matches 977; Conservative 0; Mismatches 649; Indels 6; Gaps 1;

```
Qy 25 aatggaagaagaacattagcagcgtgagagcgtccctgctgatatagtcacacccctggtc 84
Dy 7 aatggaagaagaacattatcattgacccgtgacccatttccctggtcgtgagcagc 66
Qy 85 ggcaggacaacaattatcaccaatcatctgtataaaatttgcatctcttctctg-----aagc 138
Dy 67 ggcaggacaacagatgtttacgcattatctcgttatgcagatatttcaggatgcagc 126
Qy 139 aataatgcagtcctacacaaatgaagtaataatcatatgctcaaatatttgaaacacagctg 198
Dy 127 attgacaaatgctcatcaaaagaagaatgtttatatgaagagtttttaaaattgtcgtg 186
Qy 199 ccgcttagctgttagtagaaccaatagccttggaatgaacaaatgttggtggtgatg 258
Dy 187 tcgtttagcgaagaagttttaaaagtagatgaataaaacaaacagacacaatagcgggtg 246
Qy 259 cagtgaacaacaataaaacttttttaactctgctctgctgtctttatacttaggaataacc 318
Dy 247 tagcgaataatggttgcaattttcccttccctttaaattgcatcattgtatctctggaataa 306
Qy 319 agtagcaacatcaaatgatattgacagatgcagagtggaagttactggtcatttgaatatatc 378
Dy 307 tgcgcacccgttagtataataatcattgaacgtgaattatacacagctctgtgtatgt 366
Qy 379 aaacacacactatcatgttttagtccaaagaagaagcacccgcgtttattcttgagagtagcagca 438
Dy 367 aaacacacgcaataattttgttccagaataacttttcaaaaaagtagtgaatglaaaatc 426
Qy 439 aaatcctaagtttcatataaaagtcgtagttatcgatgacgtgtacgcacattaatgagcgt 498
Dy 427 taattataaaatgatgaaactattattataattagacttaaatgaagacttaggaggtta 486
Qy 499 tgaatgcgtatctaccctgttgacgttataactgacacacacacacacacacacacacacacac 558
Dy 487 tcaatgctcacaacactttatttccaaaattcccgatattaatcttgacgttaaaaaaat 546
Qy 559 tacacaaaagaattttgaccccttgaaaaaatgcattaatattatgctcatcattggaac 618
Dy 547 taacaaaattctttaaagcagacgacgtcaggttgctgtggtgaatgtttctctggtac 606
Qy 619 aactggaattgctaaaggtgtagtagcagcagccatagagtcataactataaagattcgttca 678
Dy 607 aactggttttcgaaggagtagcgtcaactcacagaataattgttgacagattttctca 666
Qy 679 tagcaggatcccaatttatgacacgtacgtgtccacaaacacacacacacacacacacacacac 738
Dy 667 ttgcaagatcctacttttgtaacgcaattaaatccacgacagcaatttttaacggtaat 726
Qy 739 accgttccatcatcctttggaatgtttactacattattacttactttagtaggacttaaa 798
Dy 727 acccttccacacatggtttggtatgtagacacacattaggatcacttacttctggtattccg 786
Qy 799 ggttgtaattgtgaagaataattgagggcgcactttctttaaaccacacacacacacacacacac 858
Dy 787 agttgctctaatagcacacgtttgagaaaaaactatttctacaatcattacaagaattataa 846
Qy 859 aatccccactatttagtggccctccagttatggtgttttggcttaaaagccacattagt 918
Dy 847 agtggaaagtacttacttgtaccacaattaatgtgcatttttgcacaaagtgcattagt 906
Qy 919 cgatcaatcacgatttatgcagcttaacggaagtgtgactggaagagcctccttttaggaaa 978
Dy 907 tgaagaatcacgatttatgcacatttaaaagaataatgcatctggtgtggtgcacatttatcaaa 966
Qy 979 agatgcgcagagtagcagaagaggttgaaattacctggaatcatcacagaagatagg 1038
Dy 967 agaaattggggagatggttgaaaaaacgggtttaaattaaactttgtcaggcgaagggtatgg 1026
Qy 1039 attaacgaaacttgctgcgctgtaattgattacccctcataatgctgtgaaaacaggttc 1098
```

```
Dy 1027 attaacgaaacccacttgcgtgtgttttaattacacccgacacactgacgctcagaccggggtc 1086
Qy 1099 aactgggaagacccttgcatacatataaagctaaagctaaagcttttagataacgcgtactggaagc 1158
Dy 1087 aactgggtaaatagtagccatttccacgtcttaagctgtcgaatgctcgtacacacaggaataa 1146
Qy 1159 gctaggaccagaggaagagcggaataatgcttttcaaaagtgaataatgattagaaaggata 1218
Dy 1147 tttggggcacaatgaacctggaatgttattttaaagcgacatgataatgaagaagtta 1206
Qy 1219 ttacaacaactcgaagcaactattgatactactatgacaaagatggttggtcctcattctg 1278
Dy 1207 ttataataatgaagaagctactaaagcaattattacaagaacggatggttgcgtcctcgg 1266
Qy 1279 agatattggtattacgcacgaagatggaatttcttttagttgatcgtatgaagaact 1338
Dy 1267 tgatattgcttattatgacaatgatgccattttttatatatttggaacagcgtgaagtcat 1326
Qy 1339 tattaaatacagaaggatatacaggttgccgtgctgaaacgctggaataatcgtcttttacaaca 1398
Dy 1327 aattaaaataaagtttatcaggttgccactgctgaatgtgagggaatactactttacaaca 1386
Qy 1399 tccaagtattcgtgacggtgttactgaggttccggaacgaatttgggtggaacattacc 1458
Dy 1387 tccgtatattgttgcgcggttactggtataccggtatgaagcgcggcgaggttcc 1446
Qy 1459 tgcgtctgtgtgtttagaatactggaagcgtgctgactgaaaggaagtccaagatt 1518
Dy 1447 agctgcaggtgtttagtacagactggaaaatatataacagaacaaatcgtcacaaaattt 1506
Qy 1519 tattgcagcacaagctactcctcaacaaagcatcttcgagcggtgctgtattttagacag 1578
Dy 1507 tgttccagtcagtttcaacagccaaatggtactgctggtgggtgaaaatttttggatga 1566
Qy 1579 tattccgaaagccctactggaacactcatcagaagaagcgtccgagaataattttgccca 1638
Dy 1567 aatcccaaaaggaactgaactgaaaatttgacagaaaagtgttaagacaaatgttgaaaa 1626
Qy 1639 gcgagcaccacaa 1650
Dy 1627 acacaaatcctaa 1638

RESULT 6
AAS00896
ID AAS00896 standard; cDNA; 1642 BP.
XX
AC AAS00896;
XX
DT 04-JUL-2001 (first entry)
XX
DE Beetle cDNA encoding luciferase mutant LucPpe2 [T249M].
XX
KW Beetle; Luciferase; thermostability; recursive mutagenesis;
XX ATP detection; luminescence; mutant; LucPpe2; T249M; ss.
XX
OS Photuris pennsylvanica.
XX
FH Key Location/Qualifiers
CDS 8..1642
FT /*tag= a
FT /product= "LucPpe2 [T249M]"
FT /partial
FT /note= "No stop codon"
FT /transl_except= (pos:1202..1204,aa:Gly)
FT mutation replace (753,C)
XX
XX WO200120002-A1.
XX
XX 22-MAR-2001.
```

PF 22-DEC-1999: 99WO-US30925.

XX

PR 15-SEP-1999; 99US-0396154.

XX
DA / BROM-1 BROMCA CORR

FA (PROM-) PROMEGA CORP.
XX

PI Wood KV, Hall MP, Gruber M;

XX

DR WPI; 2001-244809/25.

DR P-PSDB; AA001230.
XX

Mutant luciferases from Photu

PT plagiothalamus having incre

PT to inhibitors and/or enhanced

PT detecting AIP -
XX

PS Claim 11; Fig 44; 260pp; Enql

XX

CC The sequence encodes a Beetle

CC luciferases from photuris pen
CC are created by recruiting muta

are created by recursive mutation, e.g. increased the

CC inhibitors and/or enhanced en-

CC wild-type enzyme. The mutant

for labeling a molecule, as g

solid surface, to produce hybrid reactions, or for creating junctions.

CC have increased thermostability reactions, or for creating

CC altering its activity and imp

CC assays.

XX
05
Sequence 1512 BP: 550 A: 255

```
Query Match      32.4%; Score 571.6; DB 22; Length 1642;
Best Local Similarity 59.9%; Pred. No. 9.5e-102;
Matches 977; Conservative 0; Mismatches 649; Indels 6; Gaps 1;
```

Qy 25 aatggaagaaacattagqcatqqagagcctcctgataatagtcctgqctc 84

[illegible]

Db 7 aatggaagataaaatattttatatggacctgaaccattttatcccttgctgatgggac 66

[illegible][illegible]

Db 67 ggctggagaacagatgttttacgcattatctcgttatgcagatatttcaggatgcatagc 126

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 139 aataatcgatgctcatacaagaatgaagtaatatcatatgctcaaatatttgaaaccagctg 198

Db 127 attgacaaatgcctcatcacaaagaaatgttttatatcgaagagttttttaaatttgcctg 186

Qy 199 ccgcttagctgttagtagaacaatatggcttgaatgaaacaatgttgtgggtgatg 258

[illegible]

187 ccgcccagcggagggcccccaaaacgacacaaatagcgggtgtg 246

Qy 259 caqtgaaacaataataaacttttttaatacctctgccttgcctgcttttatacttaggaatacc 318

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 247 tagcaaaatgggttgcaattttccttccttaattgcattgtatcttggaataat 306

100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098

WT ATGCTCAATCCAAATGATCATGAGTGAACCTGTCATTGGATATC 378

QY 1639 gcgagcaccaaa 1650
 Db 1627 acacaaatctaa 1638

RESULT 7

AAX32847
 ID AAX32847 standard; DNA; 1639 BP.

XX AC AAX32847;
 XX XX

DT 28-JUN-1999 (first entry)
 XX XX

XX Mutant luciferase encoding DNA.
 DE XX

XX Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;
 KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.
 XX XX

OS Photuris pennsylvanica.
 OS Synthetic.

XX W09914336-A2.
 PN XX

PD 25-MAR-1999.
 XX XX

PF 18-SEP-1998; 98WO-US19494.
 XX XX

PR 19-SEP-1997; 97US-0059379.
 XX XX

XX (PROM-) PROMEGA CORP.
 XX XX

XX Hall MP, Wood KV;
 PI XX

DR WPI; 1999-229538/19.
 XX XX

DR P-PSDB; AAY06854.
 XX XX

PT New mutant luciferase enzymes with increased stability
 XX XX

PS Claim 16; Page 81; 167pp; English.
 XX XX

CC The invention relates to a mutant luciferase enzyme that have increased
 CC thermostability. The mutant luciferase is obtained by mutating a
 CC polynucleotide sequence encoding a beetle luciferase and selecting
 CC expression products having increased thermostability. The new
 CC luciferases can be used in ATP assays, as luminescent labels for nucleic
 CC acids, proteins or other macromolecules, as genetic reporters, in enzyme
 CC immobilisation, as hybrid proteins, in high temperature reactors, and
 CC in luminescent solution. The improved thermostability allows storage of
 CC luciferases without altering its activity, and improves reproducibility
 CC and accuracy of assays using the new luciferases. The novel luciferases
 CC can also have improved brightness of luminescence, substrate utilisation
 CC and luminescence signal. Sequences AAX32843-856 represent DNA molecules
 CC encoding the mutant luciferase enzymes.
 XX XX

SQ Sequence 1639 BP; 559 A; 254 C; 324 G; 502 T; 0 other;

Query Match 32.3%; Score 570; DB 20; Length 1639;
 Best Local Similarity 60.0%; Pred. No. 1.9e-101;
 Matches 970; Conservative 0; Mismatches 640; Indels 6; Gaps 1;

QY 25 aatggaagaagaacacattagcgagcgctcgtcgatagatagccatcctggctc 84
 Db 7 aatggaagataaaaattttatggacccgtgaaccattttatcccttggctgagc 66
 QY 85 ggcaggacaacattatcaccattggtatataaaattgcatctttctctg-----aagc 138
 Db 67 ggcctggagaacagatgtttgcgcattatctcgttatgcagataattcaggatgcagc 126
 QY 139 aataatcgatgctcatacaaatgaagtaataatcatatgctcaaaatttgaaccagctg 198
 Db 127 attgacaaatgctcatacaaaagaaaattgttttatatgaagatttttaaaattgctg 186

QY 199 ccgttagctgttagtagaacaataatggttggaatgaaacaaatgttgggtgtgtg 258
 Db 187 tcgttagcggaagtgttaaaagtagtggttaaaacaaacgacacaatgctgtgtg 246
 QY 259 cagtgaacaaacataaaactttttaactcctgtcctgtcctgtcttatacttaggaatacc 318
 Db 247 tagcgaataatggttgcgaatttttctcctataattgcatactgttatcttggaaataat 306
 QY 319 agtagcaacatcaaatgatatgtacacagatggagagttaactggtcatttggatatact 378
 Db 307 tgcgacacctgttagtgataaatacatgaactgaactgaataataacacagctcttggatagt 366
 QY 379 aaacccaactatcattgttttagtccaagaagcactccgcttatttctgagagtagacgca 438
 Db 367 aaacccaactatcattgttttagtccaagaacttcttcaaaagtagtgaatgtaaatc 426
 QY 439 aaactaagtttcaataaaagtcgttagttatcgatagcgttacgacattaaatggcgt 498
 Db 427 taataataaataatgtagaaactattattattagaaacttaataagaacttaggaggtta 486
 QY 499 tgaatgcglatctacctttgttgcaactgttatctgaccacacacttgcatttgcatttgcatt 558
 Db 487 tcaatgcctcaacaaactttattctcaaaattcogataatttaataatcttgcgtlaaaaaatt 546
 QY 559 tacacaaaagaattttgatcccttgaaaaaactgcattaaattatgtatcatcatctgggaac 618
 Db 547 taacattattcttttaactcgagcgtcaggttgcgttggtaatttcttctcgtgtac 606
 QY 619 aactggattgcctaaggtgttagtactgagcctagaagagcttaactataaagattcgttca 678
 Db 607 aactgggttttcgaaggagtcagttaactcaagaataattgttgcacgattttctca 666
 QY 679 tagcgggatcccatatttgaactcgtcgtgttcacaaacacatcaatttcttctttagt 738
 Db 667 tgcaaaagatcctactcttctgtgaacgcaatttaataccacagacgcaattttaaocggtta 726
 QY 739 accgttccatcatgctcttggaaattttactacattacttacttactttagtagtaggactaa 798
 Db 727 acctttccacctggtttgtgtgatgaccacatagtagacttacttcttctgtgtgtcgtg 786
 QY 799 ggttgttaattgtgaagaatttggggcgcaactttcttcaaaacacacatacagaatacaaa 858
 Db 787 agtgttctaataatgcacggtttgaagaaaaactatttcaatactattacaagataataa 846
 QY 859 aatcccaactattgttagtggccctccagttatggtgtgttttggctaaagccccattagt 918
 Db 847 agtggaaagtacttacttactgtaccacaaacttaagtgcatttttgcacaaagtgcattagt 906
 QY 919 cgatcaatcacgattttatcgagcttaacggaagttcgtcactgagagagcctctttaggaaa 978
 Db 907 tgaagagcacgatttactgcgactttaaagaaattgcactctgtgtggcgacacctttatcaaa 966
 QY 979 agatgtcgcagaagcagtagcaaaagaggtgaaatttacttgaatacacaaggagataagg 1038
 Db 967 agaaattggggagatggtgaaaaaacggttttaataacttctgcaggcgaaggtatgg 1026
 QY 1039 attaaactgaacttgcgtgcgtgttaattgattaccctccataatgtctgaaacacaggttc 1098
 Db 1027 attaacagaacacactctcggtgttttaattacacgcgaacaaatgacgtcagaccgggatac 1086
 QY 1099 aactggaagacaccttgcctacatacattaaagctaaaagtttttagataacgctacttggaaaggc 1158
 Db 1087 aactggtgaaatagtagcattttccacgtgtttaaagtttgcgactctacacacgggaaaaat 1146
 QY 1159 gctaggaccaggagaagagggcggaataatgctttcaagtgaaatgattatgaaaggata 1218
 Db 1147 tttggggcccaaatgaaactggagaattgtatttttaaaaggcagacagataatgaaagggtta 1206
 QY 1219 ttacaacaaatccggaagcaactattgatactatttgcacaaagatggttggcttctcctgg 1278
 Db 1207 ttataataatgaagagctactaaagcaattatttaacaaagcaggtggttgcgtcctgg 1266
 QY 1279 agatatgtggattattacgacggaagatggaaaatttctttattagttgatcgattgaaagaact 1338

Db 1267 tgaattgctcatatgacaatgatggccattttatatattgtgacagctgaagtcatt 1326
Qy 1339 tatataacacagggatcatcaggttgccctgctgaaactggaaatcgtctttacaaaca 1398
Db 1327 aattataataaaggattatcaggttgccctgctgaaactggaaatcgtctttacaaaca 1386
Qy 1399 tccaaagtattgctgacggttgctgactggaagttcccgacgaatttggcgacaattacc 1458
Db 1387 tccgtatatattgtagtcggtgctgacgtgacgtgacgtgacgtgacgtgacgtgacgt 1446
Qy 1459 tgcgtctgtattgttagaataatggcaagacgctgactgactgactgactgactgactgact 1518
Db 1447 agctgcaggtgtttagtacagactggaaatacttaacagcaacaaatcgtacaaattt 1506
Qy 1519 tatcgacgacaaagtcactccaaagcattcttcgagcgttgcgagcgtgtcgtattgtagacag 1578
Db 1507 tgtttccagtcagatttcaacagccaaatggctacgtggtggtgnaatttttggatga 1566
Qy 1579 tattccgaaagccctacttgaaactcatcagaaaggagctccgagaaatatttg 1634
Db 1567 aattcccaaggatcaactggaaatgtgacagaaatgttaagacaaatgtttg 1622

RESULT 8

ID AAX32844
XX AAX32844 standard; DNA; 1639 BP.

AC AAX32844;

DT 28-JUN-1999 (first entry)

DE Mutant luciferase encoding DNA.

KW Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.

OS Photuris pennsylvanica.

OS Synthetic.

XX WO9914336-A2.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19494.

XX 19-SEP-1997; 97US-0059379.

XX (PROM-) PROMEGA CORP.

XX Hall MP, Wood KV;

XX WPI; 1999-229538/19.

XX P-PSDB; AAY06851.

XX New mutant luciferase enzymes with increased stability

XX Claim 16; Page 79; 167pp; English.

XX The invention relates to a mutant luciferase enzyme that have increased
XX thermostability. The mutant luciferase is obtained by mutating a
XX polynucleotide sequence encoding a beetle luciferase and selecting
XX expression products having increased thermostability. The new
XX luciferases can be used in ATP assays, as luminescent labels for nucleic
XX acids, proteins or other macromolecules, as genetic reporters, in enzyme
XX immobilisation, as hybrid proteins, in high temperature reactors, and
XX in luminescent solution. The improved thermostability allows storage of
XX luciferases without altering its activity, and improves reproducibility
XX and accuracy of assays using the new luciferases. The novel luciferases
XX can also have improved brightness of luminescence, substrate utilisation
XX and luminescence signal. Sequences AAX32843-856 represent DNA molecules
XX encoding the mutant luciferase enzymes.

Qy 25 aatggaagaagaacattagcgatggagagcgctcctcgtgatatagtccatcctggctc 84
Db 7 aatggaagataaataattttatatatggacctgaacattttcctcctggctgagggagc 66
Qy 85 ggcgggacacaaattataccatcattgtataaaatttgcattctttctctg-----aagc 138
Db 67 ggcgtggagacagagtgtttacgcattatcgtttatgcagataatttcaggatgcatagc 126
Qy 139 aataatcgatcctcatcacaaatgaagtaataatcatatctcaaatatttgaaccagctg 198
Db 127 attgacaatgctcatatacaaaagaaaatgtttatatgaagagtgttaaaattgtcgtg 186
Qy 199 ccgcttagctgttagatagacaataatgtgttgaatgaacaacaatgttgggtgtatg 258
Db 187 tcgttagcggaaagtgtttaaagatagattaaacaaacacacacacacacacacacacac 246
Qy 259 cagtgaacacaaatataaaacttttttaattcctcgtccttgcgtgttttatacttaggaatacc 318
Db 247 tagcgaaatggtttgcatttttctcctcctataatgtcatcattgtatcttggataatc 306
Qy 319 agtagcaacatcaaatgatagtacacagatggagaggttaactgctcatttgaatatatc 378
Db 307 tgcagacctgttagtgataaatacatcattgaacgtgaattatacacagctccttggattgt 366
Qy 379 aaaaaccaactatcatgttttagttcaaaagaagcactccgcgttatcttgagagcacgca 438
Db 367 aaaaaccagcaataatttttgcctccaaagaatacttttcaaaagctactgaaatgtaaaatc 426
Qy 439 aaatcttagttcattaaaaagtcgtagtattatogatagcattgtacgacattaaatggcgt 498
Db 427 taaattaaaatatgtagaactattattatttagaactaaatgaagacttagagggtta 486
Qy 499 tgaatgcgtatcctacctttgtgcacgttatctactgacacacacacacctttgacatctgcaatt 558
Db 487 tcaatgcctcaacactttattctcaaaattcogataattaatctgggacgttaaaaaaatt 546
Qy 559 tacacaaaagattttgatcccttgaaaaaatcgcatataatgtcatcatctcgtggaac 618
Db 547 taaccataattcttttaactgagacgatacaggttgcgttgtaagtgtttcttctcgtgac 606
Qy 619 aactggattgcttaagggttagtactgagccatagaaagtctaaactataagattcgttca 678
Db 607 aactggtgtttcgaaggagtcgtactaaactcaacaagataattgttgacgattttctca 666
Qy 679 tagcagggatcccatattatggcactcgtacggttcccaacaacatcaattcttctcctagt 738
Db 667 tgcaaaagatcctactcttttggtaacgcaatttaataccacgacagcaattttaaocggtaat 726
Qy 739 accgttccatcatgccttttgaatgtttactacattatcttacttcttagtagaagctcaa 798
Db 727 acccttccacatggtttttggtagatgacacacattaggatactttacttttggattccg 786
Qy 799 ggttgaattgtgagaataatttgaggcgccacttttcttaaaaaaccatacagaattacca 858
Db 787 agtgttctaatagcacacgcttgaagaanaaactatttacaactattacaagattataa 846
Qy 859 aatccccactatttagtggccctccagttatggtgttttttggctaaaaagcccaattagt 918
Db 847 agtggaaagtacttactgttaccacacattaatggcatttttttgcanaaagtgcattagt 906
Qy 919 cgataacacagatttatcgcgcttaaacgaaagtgcctactgagagagctcctcttaggaaa 978
Db 907 tgaagaagacagatttatcgcacttaaaagaanaattgcactcgtggtgcgcacctttatcaa 966
Qy 979 agatgtcgcagaagcagtagcaaaagaggttgaatttaactcctggaaatcacacaagggatg 1038
Db 979 agatgtcgcagaagcagtagcaaaagaggttgaatttaactcctggaaatcacacaagggatg 1038

Db 967 agaaattggggagatggtgaaaaaacgggttttaaatataaacttttgtcaggcaagggtatgg 1026
QY 1039 attaaactgaactgtgcgtgctgaatgaattaccctcataatgctgtgaaacacaggttc 1098
Db 1027 attaacagaacaccttcggtgttttaattacaccgaacaatgacgtcagacggggtac 1086
QY 1099 aactgggaagaccttgcatacatataaagctaaagtatttagataaacgtctactggaagcc 1158
Db 1087 aactcgttaaatagtaaccttaccgtgtttaaagtgtcgtacctcaacagagaaaaat 1146
QY 1159 gctagaccagggaagagcgcaatatctttcaaatgaaatgattatgaagagata 1218
Db 1147 ttggggccaaatgaactggagatgtattttaaaggcgacatgataatgaaggtta 1206
QY 1219 ttacaacaatccggaagaactattgatactattgacaaagatgggttgccttcatctgg 1278
Db 1207 ttataataatgaagaagctactaaagaacaattatttaacaagaagcggatgggtgcgctctgg 1266
QY 1279 agatattggattattacagcaagatgaaatctttcttatagttgatcgattgaaagaact 1338
Db 1267 tgatattgcttattatgacaatgatggccatttttataattgtagcagggtgaagtcat 1326
QY 1339 tattaaatacaagggtatcagggttgcgctgctgaaactgaaatactgcttttacaaca 1398
Db 1327 aattaataataaaggttatcagggtgcacctgtgaaatgagggaataactcttacaaca 1386
QY 1399 tcaaggtattgtagatcggggttactgaggtccgggaagaaatttggtagacaattacc 1458
Db 1387 tccgtatattgtgtagccgggttactggtataccggtatgacccggcgagcttcc 1446
QY 1459 tgcgtctgtgtgtgttagaattctgcaagacgtcactgaaagaaggaagttcaagattt 1518
Db 1447 agctgcagggtgtgtgtagacagactggaaaatacttaaaacgaaacaaatcgacaaaatt 1506
QY 1519 tattgcagcaagaactcactcaacaaagcatcttcgaggcggtgtcgtattttagacag 1578
Db 1507 tgtttccagtcagatttcaacagcagcaaatggtcactggtggtggaatttttgatga 1566
QY 1579 tattccgaagggccctactgaaactcaatcagaagaaggtcccgaaataatttg 1634
Db 1567 aattcccaaggatcaactggaaaattgacagaaaagtgttaagacaaaattgtttg 1622

RESULT 9
AAX32843
ID AAX32843 standard; DNA; 1639 BP.
XX AC AAX32843;
DT 28-JUN-1999 (first entry)
XX Mutant luciferase encoding DNA.
DE Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.
XX OS Photuris pennsylvanica.
OS Synthetic.
XX MO9914336-A2.
XX PN 25-MAR-1999.
XX PD 18-SEP-1998; 98WO-US19494.
XX PF 19-SEP-1997; 97US-0059379.
XX PR (PROM-) PROMEGA CORP.
XX PA Hall MP, Wood KV;
XX PI WPI; 1999-229538/19.
XX DR P-PSDB; AAY06850.
DR

XX New mutant luciferase enzymes with increased stability
PT Claim 16; Page 79; 167pp; English.
XX The invention relates to a mutant luciferase enzyme that have increased
CC thermostability. The mutant luciferase is obtained by mutating a
CC polynucleotide sequence encoding a beetle luciferase and selecting
CC expression products having increased thermostability. The new
CC luciferases can be used in ATP assays, as luminescent labels for nucleic
CC acids, proteins or other macromolecules, as genetic reporters, in enzyme
CC immobilisation, as hybrid proteins, in high temperature reactors, and
CC in luminescent solution. The improved thermostability allows storage of
CC luciferases without altering its activity, and improves reproducibility
CC and accuracy of assays using the new luciferases. The novel luciferases
CC can also have improved brightness of luminescence, substrate utilisation
CC and luminescence signal. Sequences AAX32843-856 represent DNA molecules
CC encoding the mutant luciferase enzymes.
XX SQ Sequence 1639 BP; 558 A; 255 C; 323 G; 503 T; 0 other;

Query Match 32.1%; Score 566.8; DB 20; Length 1639;
Best Local Similarity 59.9%; Pred. No. 8.1e-101;
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

QY 25 aatggaagaagaacatttaggcgtgagagcgtcctcgtgatagatagtcctcctggtc 84
Db 7 aatggaagataaaaattttatattgacacctgaaccattttatcccttgcgtgagggac 66
QY 85 ggcaggacaacaattataccaatcattgtataaaaatttgcatcttctctg-----aagc 138
Db 67 ggcaggagaacagatgttttacgcattatctcgttatgcagataatttcaggatgcatagc 126
QY 139 aataatcgatcctcatacaaatgaagtaataatcatatctcctcaaatatttgaacacagctg 198
Db 127 atgacaaatgctatacaaaagaataatgttttatagaagatttttaaaattgctgtg 186
QY 199 ccggttagctgttagtagaacaataatggtctgaatgaagaacaataatgttg-gggtgtatg 258
Db 187 tcgtttagcggaaagtgtttaaagatggtgattaaacaaacagacacaatagcggtgtg 246
QY 259 cagtgaacacaataaaacttttttaactcgtcctcgtcttcttatacttaggaataacc 318
Db 247 tagcgaataatggttgcgaatttttccctccctataatgcatcattgtatctctggaataat 306
QY 319 agtagcaacatcaaatgatatgtacacagatggagagtgtaactggctcatttgaatataatc 378
Db 307 tgcagcactgttagtgataaatacattgaacgtgaaataatacacagcttctgttatgt 366
QY 379 aaacccaactatcatgttttagttccaagaagaagcactcccgcttattctgtgaggtacagca 438
Db 367 aaacccagcataaatttttgcctcagaatacacttttcaaaaagtactgaatgaaaaatc 426
QY 439 aaactcaagtttcatttaaaaaagtcgttagttatgcagatgacgtacgacattaaatgcgt 498
Db 427 taaatataaataatgtagaaactattattattagacttaataatgaagacttaggggtta 486
QY 499 tgaatgcgtatctaccttctgtgcacgttatactagcaccacacaccttgcattcattgcatt 558
Db 487 tcaatgcctcaacaactttattctcaaaaattccgataattaatcttgacgtaaaaaaatt 546
QY 559 tacacaaaagaattttgatcccttgaaaaaatgcattaaattatgtcatcatctcgtggaac 618
Db 547 taacataattctttaaactcgagacgatcagggttgcgttggtaagtgttttctctcgttac 606
QY 619 aactggattgcctaaggtgtgtagtactagccatagaaggtctaaactataaagattcgttca 678
Db 607 aactgggttttcgaaggagtcactgactcaactcaagaataattgttgacagatttctct 666
QY 679 tagcagggatcccaattttatggcactcgtacgttcccaacacacataattcttctcctagt 738
Db 667 tgcaaaagatcctactcttttggtaacgcgaatttaataccacgacagcaattttaaagggtaat 726

```
Qy 739 accgttccatcatgcccttctggaattgtttactacattatcttacttcttagtaggaactaa 798
Db 727 acctttccaccatggcttctgtgatgacacacattggatacttctacttctgtgtccog 786
Qy 799 ggttgtaattgtgaagaatttggagcgcaactcttcttaaaacccatcacagaattacaa 858
Db 787 agtggctctaattgcacagcttttgagaaaaaactatttctacaatcattacaagattataa 846
Qy 859 aatccccactattgtatggtccctccagttatgtgttcttttggctaaaagcccatagt 918
Db 847 agtggaaagtacttctacttgaccacaattaatgtgcatttcttgcacaaagtgcattagt 906
Qy 919 cgatacatagatttatcgactgacttaacggaagtgtctactctggagagactctcttagaaa 978
Db 907 tgaagtagcagatttctcgcattcaagaataatgcactctgtgtggtgcacactttatcaaa 966
Qy 979 agatgcgcagagcugtagcagaagaggttgaaattaccctggaatcacatcacagaatatgg 1038
Db 967 agaaattggggagatggtgaaaaaacggttttaaatataaactttgtcaggcaagggtatgg 1026
Qy 1039 attaaactgaaactgtgcgtgctgtaagtattaccctccatcaataatgctgtgaaaaacaggttc 1098
Db 1027 attaacgaaacacttcggtgtgttttaattacaccgcaacaatgacgtcagaccggatc 1086
Qy 1099 aactgggaagacccttgcatacatataaagctaaagtgttttagataacgcgtactctgggaaggc 1158
Db 1087 aactggtaaaatagatcatttcacgtgtttaaagtgtcgtactcctcaacaggaataat 1146
Qy 1159 gctagggaccggagaaagagcgaaatatgcttccaaagtgaatgatattgaagaagata 1218
Db 1147 ttggggccaaatgaactggagaatgtgtattttaaaggcgacatgataaagaaggtta 1206
Qy 1219 ttacaaacaatccggaagacatattgtactactatgcacaaagataggttggtcttcatttgg 1278
Db 1207 ttataataatgaagaagctactaaagcaattatttaacaaagacggatggttgctctctgg 1266
Qy 1279 agatattggattatcacacgaagatggaaaatttctttatagttgatcgatgaaagaact 1338
Db 1267 tgatatgtctattatgacaaatgatggccatttttatatttggacagggctgaagtcatt 1326
Qy 1339 tattaatacagaagggatcacaggttgcgctgctgaaactggaataatctgttttacaaaca 1398
Db 1327 aattaaataaaggattacaggttgccactgctgaaattgaggggaataactctttacaaaca 1386
Qy 1399 tccaagtattgctacgcgggtgtactgaggttcgcgacgaagaatttggtagacaattacc 1458
Db 1387 tccgtatattgtctgacggcgctactgggtacacggatgaagccgagcggtcc 1446
Qy 1459 tgcgtctgtgtgttagaactctggcagacgcgtgactgaaaggaaggttcgaagattt 1518
Db 1447 agctgcagggtgtgttagtacagactggaaaatatctaaacgaacaaatcgtaaaaaattt 1506
Qy 1519 tattgcgacaaactcactccaaacaaagcatctctgggggggtgcgttatgttagacag 1578
Db 1507 tgtttccagtcacagtctcaacagacaaatggctacgtggtgggtgaaatttttggatga 1566
Qy 1579 tattcccaaggccactgctggaacactcatcagaagaagagctccgagaaaattttg 1634
Db 1567 aattcccaaggatcaactggaaaaaattgacagagaagtgcttaagncanaatggtttg 1622
```

```
RESULT 10
AA32845
ID AA32845 standard; DNA; 1639 BP.
XX
AC AA32845;
XX
DT 28-JUN-1999 (first entry)
XX
DE Mutant luciferase encoding DNA.
XX
KW Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;
```

```
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.
XX Photuris pennsylvanica.
OS Synthetic.
XX WO9914336-A2.
XX 25-MAR-1999.
XX 18-SEP-1998; 98WO-US19494.
XX 19-SEP-1997; 97US-0059379.
XX (PROM-) PROMEGA CORP.
XX Hall MP, Wood KV;
XX WPI; 1999-229538/19.
XX P-PSDB; AAY06852.
XX New mutant luciferase enzymes with increased stability
XX claim 16; Page 80; 167pp; English.
XX The invention relates to a mutant luciferase enzyme that have increased
thermostability. The mutant luciferase is obtained by mutating a
polynucleotide sequence encoding a beetle luciferase and selecting a
expression products having increased thermostability. The new
luciferases can be used in ATP assays, as luminescent labels for nucleic
acids, proteins or other macromolecules, as genetic reporters, in enzyme
immobilisation, as hybrid proteins, in high temperature reactors, and
luciferases without altering its activity and improves reproducibility
and accuracy of assays using the new luciferases. The novel luciferases
can also have improved brightness of luminescence, substrate utilisation
and luminescence signal. Sequences AAX32843-856 represent DNA molecules
encoding the mutant luciferase enzymes.
XX Sequence 1639 BP; 557 A; 255 C; 323 G; 504 T; 0 other;
```

```
Query Match 32.1%; Score 566.8; DB 20; Length 1639;
Best Local Similarity 59.9%; Pred. No. 8.1e-101;
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

Qy 25 aatggaagagaaacacattaggaatggagagcgctccctgtagatagtagccatccgtgctc 84
Db 7 aatggaagataaaataatttttatatggacctgaaccattttatcccttggctgtagggac 66
Qy 85 ggcaggagcaacaattatcaccaatcatgtataaaatttgcattcttttccctg-----aagc 138
Db 67 ggcaggagaaacagatgttttacgcattctctcgttatgcagatatatttcaggatgcatagc 126
Qy 139 aataatcgatgctcatcacaaatgaagtaatatcatatgctcaaaatatttggaaacagctg 198
Db 127 attgacaaatgctcatacaaaagaaaatgttttatagaagagtttttaaaattgtcgtg 186
Qy 199 ccgtttagctgttagtagatagacaataatgcttgaagaaaaaacaatgttctgggtgtatg 258
Db 187 tcgtttagcggaaagtgttttaaaagtgatgattataaaacaaacacacaaatagcgggtgtg 246
Qy 259 cagtgaacaaacataaaacttttttaactcctgtctctgtctgttatactttaggaataacc 318
Db 247 tagcgaataatggtttgcgaatttttccctccatataatgcacattgtctcttggaaataat 306
Qy 319 agtagcaacatcaaatgatatgtacacagatggagagtgtaactggtcatttgaatatatc 378
Db 307 tgcagaccctgttagtataaatacattgaaacgtgaattaatatacacacagcttggattgt 366
Qy 379 aaaacaaactatcatgttttagttcaagaagaagcactcccgcttattcttgagagtagacga 438
Db 367 aaaacacgcgataatttttctgctccagaaacacttttccaaaaagactgaaatgataaatac 426
```


XX Mutant luciferases from Photuris pennsylvanica and Pyrophorus
PT plagiophthalmus having increased thermostability, increased resistance
PT to inhibitors and/or enhanced enzymological properties, useful for
XX detecting ATP -
XX
XX Disclosure; Fig 26; 260pp; English.
XX
XX The sequence encodes a Beetle mutant luciferase, Luc49-4G11. Mutant
CC luciferases from Photuris pennsylvanica and Pyrophorus plagiophthalmus,
CC are created by recursive mutagenesis to have one or more desired
CC properties, e.g. increased thermostability, increased resistance to
CC inhibitors and/or enhanced enzymological properties relative to a
CC wild-type enzyme. The mutant luciferases are useful for detecting ATP,
CC for labeling a molecule, as genetic reporters, for immobilisation onto a
CC solid surface, to produce hybrid proteins, for high temperature
CC reactions, or for creating luminescent solutions. The luciferase enzymes
CC have increased thermostability which allows storage of the enzyme without
CC altering its activity and improves reproducibility and accuracy of
CC assays.
XX
XX Sequence 1639 BP; 555 A; 258 C; 325 G; 501 T; 0 other;

Query Match 32.1%; Score 566.8; DB 22; Length 1639;
Best Local Similarity 59.9%; Pred. No. 8.1e-101;
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

Qy 25 antggaagaagaaacattagggatggagagcgtcctcgtgatatagtcacatcctggctc 84
Db 7 aatggaagataaaatattttatagccctggaacccattttatcccttggctgagggac 66
Qy 85 ggcaggaacacattatataccatcattgtataaaatttgacatctttctg-----aagc 138
Db 67 ggcaggaacacagatgtttgacgcatcttcgttatcagatatctcagatgcatagc 126
Qy 139 aataatcgatgctctacataaatgaagtaataatcatatgctcaaatatttgaaacacagctg 198
Db 127 attgcaaatgctctacataaatgccccctgttttatatgagaagttgttcaaaatgctg 186
Qy 199 ccgcttagctttagtatagacaataatggttggaatgaaacaaatgtgtggtgtagtg 258
Db 187 tcgttagcgaaggttttaaaagtagtgattaaacaaacacacacacacacacacacacacac 246
Qy 259 cagtgaaacacataaaacttttttaactcctgctccttctgctcttacttagaataacc 318
Db 247 tagcgaataatggttgcaatttttctcctcctataaattgcatcattgtatcttggaaat 306
Qy 319 agtagcaacatacaatgatatgtacacacagatggagagtgtaactggtcatttgaatatatc 378
Db 307 tgcagcacctgttagtataatacatatgacgtgaaattataacacacacacacacacacacac 366
Qy 379 aaac 438
Db 367 aaac 426
Qy 439 aaatcagatttccataaaagtcgttagttatcgatagacatgacacacacacacacacacacac 498
Db 427 taataataaatatgagaacacatattatataatagacatgaatgaagacttagagggtta 486
Qy 499 tgaatgcgtatctacacttctgtgacgtttactgacacacacacacacacacacacacacacac 558
Db 487 tcaatgcctcaac 546
Qy 559 tac 618
Db 547 taac 606
Qy 619 aactggattcgtcaggtgttagtactgagccatagacatagacatgaatcacaatcgaattcgtca 678
Db 607 aactgggttctcgaaggagtcagtctactcaactcaacagacatattgtgcacgattcttcca 666
Qy 679 tagcagggatcccatttatggcactgctacggttccacacacacacacacacacacacacacac 738

Db 667 tgaagaagatcctactcttctgttaacgcaataatcccaacgacagcaattttaaaggtaat 726
Qy 739 accgttccatcatgctccttggaaatgtttactacattatcttacttcttagtaggacttaa 798
Db 727 accttccaccatggttttgatgatgacacacattagatcacttcttctgtgattccg 786
Qy 799 ggttgtaattgtgaagaatttgagggcgacacttcttcttaaaacacacacacacacacacac 858
Db 787 agttgttctaagtcacacgcttgaagaaaaactatttcaaatcattacacagacttaa 846
Qy 859 aatcccccactattgtagtgccctccagttatgtgttttttggctaaagccacttagt 918
Db 847 agtggaaagtactttacttgaacacacataatgacatttttgcacaaagtgcattagt 906
Qy 919 cgatacatcattatctatcagcttaacggaagtgtcactggaagagctcctttaggaaa 978
Db 907 tgaagaatcacgattttatcgctctaaagaataatgcatctggtggtgcacactttatcaaa 966
Qy 979 agatgtcgacagaagcagtagcaaaagaggttgaataatcactggaatcattacacaggaatg 1038
Db 967 agaaattggggagatggtgaaaaaacggcttaaaacttaaaacttgcaggcgaaggatcgg 1026
Qy 1039 attaaactgaacttgcgtcgttaataatgattaccctcctataatgctgttgaaaaacaggttc 1098
Db 1027 attaaacagaacacacttcgctgttttaataatcacccgacaacatgactcagacccgggac 1086
Qy 1099 aactggaaagaccccttgcctacatttaagcttaagcttttagataacgctactggaagac 1158
Db 1087 aactggttaaatagtagtaccatttcacgctgttaaggttgcactcactacacacaggaanaat 1146
Qy 1159 gctaggaacccggaggaagcgcgaataatgcttcaagagtaaatgattatgaagaagata 1218
Db 1147 tttggggcccaatgaactgggaatgtatttttaagggcagacatgataaaagaggtta 1206
Qy 1219 ttacaacaatccggaagcaactattgatactattgacaaaagatggttggtctcattctgg 1278
Db 1207 ttataataatgaagaagctactaaagcaattattacaacaaagacgagatggttgcgtctcgg 1266
Qy 1279 agatatggatattacacgacgaagatggaattttcttattagttgatcgtattgaaagaaact 1338
Db 1267 tgatatgtcttatatgacaatgagtgccattttattattgtgacacaggtcgaagtcatt 1326
Qy 1339 tatataacacaaagggatcacggttgcctcgtgaactggaataatcgtctttacaasca 1398
Db 1327 aattaaaataaaaggttatcaggttgcacctgctggaataatbtgggggaataactcttacaaca 1386
Qy 1399 tccaagtattcgtgatgcgggtgttacttggaggttccggacgaatttggatggacaaatcacc 1458
Db 1387 tccgtatattgtgatgcggcggttacttggatataccggatgaagcgcggcgagcttccc 1446
Qy 1459 tgcgtctgtgttctgtttagaatctggcaagacgctgactgaaagaggaagttcaagaattt 1518
Db 1447 agctgcaggtgttctgttagcacagactggaaaataatactaaacacacacacacacacac 1506
Qy 1519 tatgacgacaaagtccactccaacaaagcatcttcggagcgggtgcgtatttcttagacag 1578
Db 1507 tgtttccagtcagaagtctcaacagccaaatggctacgtggttgggttgaatttttggagaga 1566
Qy 1579 tattccgaagagccctactctggaaaactcatcagaaaaggagctcccgagaaatattttg 1634
Db 1567 aattcccaagagatcaactcggaaaaattgacagaaaagtgttaagacaaaattgttg 1622

RESULT 12
AA32846
ID AA32846 standard; DNA; 1639 BP.
XX
AC AA32846;
XX
DT 28-JUN-1999 (first entry)
XX
DE Mutant luciferase encoding DNA.


```
QY 679 tagcaggatcccatcttgccactcgtagcgtttcccaacaacatcaattcttctttagt 738
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 667 tgaagaagatcctacttttggttaacgaattatcccaacagacgaatttttaacggtaat 726
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 739 acogttccatcgtcctttggaattttactatcatcttacttctttagtaggacttaa 798
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 727 accttccacacagtttttggttagtagcaccacattaggatacttacttctgtgattccg 786
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 799 ggttgtaattgaagaaattgaggcgacacttttttaaaaccatcacagaattacaa 858
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 787 agttgttctaatgcacacgtttgaagaaaactattctcaaatcattacaagattataa 846
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 859 aatcccactatttagtggccctccagttatgtgtgttttggctaaaagcccatagt 918
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 847 agtggaaagtactttacttaccacaattaaatggcatttttgcataaaagtgcattagt 906
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 919 cgatcaatacagatttatcgacttaacgggaagtgtctactggagagctctcttaggaaa 978
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 907 tgaagaagtagatttatcgcaacttaaaagaaaattgcatctgtgtggcgcaactttatcaa 966
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 979 agatgtcgcagaagcagtagcaaaagaggttgaaattacctgggaatcatacagaattatgg 1038
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 967 agaaattggggagatgggtgaaaaaacggtttaaattaaactttgtcaggcaagggtatgg 1026
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1039 attaactgaaactgtgcgtgttaagtattaccctccataaatgctgtgaaaacaggttc 1098
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1027 attaacagaaccactcggcgtgttttaattacacgaacaattacgtcagacggggtac 1086
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1099 aactgggaagacctgcccatacatataagctaaagtgttttagataaacgtactctgggaagggc 1158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1087 aactggttaaatagtagtaccatttccgctgttaaggttgcgtctcctacacaggaaaaat 1146
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1159 gctaggaccaggagaagagcgaaatattgctttcaagtgaaatgatttgaagaagata 1218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1147 ttggggccaaatgaactgagaaattgtattttaagggcgacatgataatgaagatta 1206
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1219 ttacaacaatccgggaagcaactattgatactattgacaaagatggttggtcttctctcg 1278
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1207 ttataataagaagaagactactaaagaacaattattacaagaagcggatggttgccgtctcg 1266
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1279 agatattggaattatcacagcaagatggaattttcttatagttgattgacgattgaaagaact 1338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1267 tgattgtcttattatgacaatgatggccatttttatattgtgacagggctggaagtcatt 1326
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1339 tattaaatacaagggatatacaggttgccctgctggaactggaaatctgcttttacaaca 1398
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1327 aattcaatacaaaaggttatcaggttgccacctgctggaattgaggggaatactcttacaaca 1386
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1399 tccaagtattgctgactcgggtgttacttgagttcccgacgaattttggtgacaaattacc 1458
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1387 tccgtattattgtatgcccggttactgttataccggatgaagccggcgaggttcc 1446
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1459 tgcgtctgtgtgtgttagaatctgcgaagacgctgactgaaagaagagttcaagattt 1518
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1447 agctgcaggtgtgttagtagacagtggaataatctaaacgaacaaatcgtaacaaattt 1506
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1519 tattgcagcaacagctcactcaacaagacatcttcgagggcgtgtctgtattttagacag 1578
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1507 tgttccagtcgaatttcaacagccaaatggctacggtggtgggtgaaatttttggatga 1566
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1579 tattccgaagggcctactgaaaacctcatcagaaagagctcccgagaaattatttg 1634
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1567 aattcccaaggatcaactggaaataattgacagaaaaagtgttaagacaaatgtttg 1622
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 14

AAS00887

ID AAS00887 standard; cDNA; 1639 BP.

XX AC AAS00887;

XX DT 04-JUL-2001 (first entry)

XX

DE Beetle cDNA encoding luciferase mutant Luc49-OG12.
XX
KW Beetle; Luciferase; thermostability; recursive mutagenesis;
RW ATP detection; luminescence; mutant; Luc49-OG12; ss.
OS Photuris pennsylvanica.
OS Synthetic.

XX Key Location/Qualifiers
FT CDS 8..1639
FT /*tag= a
FT /product= "Luc49-OG12"
FT /partial
FT /note= "No stop codon"
FT trans_except= (pos:149..154,aa:Glu-Asn)
FT trans_except= (pos:665..667,aa:Leu)
FT trans_except= (pos:887..889,aa:Phe)
FT trans_except= (pos:1241..1243,aa:Thr)
FT mutation replace (150..153,AAA)
FT /*tag= b
FT mutation replace (278,T)
FT /*tag= d
FT mutation replace (388,T)
FT /*tag= e
FT mutation replace (554,A)
FT /*tag= f
FT mutation replace (665,C)
FT /*tag= g
FT mutation replace (668..670,TGC)
FT /*tag= h
FT mutation replace (792,C)
FT /*tag= i
FT mutation replace (887,T)
FT /*tag= j
FT mutation replace (1064,G)
FT /*tag= k
FT mutation replace (1068,C)
FT /*tag= l
FT mutation replace (1202,A)
FT /*tag= m
FT mutation replace (1632..1635,AATC)
FT /*tag= n
FT mutation replace (1637..1639,AAGCTG)
FT /*tag= o

WO200120002-A1.

XX 22-MAR-2001.

XX 22-DEC-1999; 99WO-US30925.

XX 15-SEP-1999; 99US-0396154.

XX (PROM-) PROMEGA CORP.

XX Wood KV, Hall MP, Gruber M;

XX WPI; 2001-244809/25.

XX P-PSDB; AAU01221.

XX Mutant luciferases from Photuris pennsylvanica and Pyrophorus
XX plagiophthalmus having increased thermostability, increased resistance
XX to inhibitors and/or enhanced enzymological properties, useful for
XX detecting ATP.

XX Disclosure; Fig 24; 260pp; English.

XX The sequence encodes a Beetle mutant luciferase, Luc49-OG12. Mutant
XX luciferases from Photuris pennsylvanica and Pyrophorus plagiophthalmus,
XX are created by recursive mutagenesis to have one or more desired
XX properties, e.g. increased thermostability, increased resistance to
XX inhibitors and/or enhanced enzymological properties relative to a
XX wild-type enzyme. The mutant luciferases are useful for detecting ATP.

Db 1147 ttctggggccaaatgaactggagaattgtatttttaaaggcgacatgataatgaagggta 1206
QY 1219 ttacaacaatccggaagcaactatgtactatttgacaagaatggttgcttcattctgg 1278
Db 1207 ttataataatgaagagctactactaagaatattatacaaaagacggatggttgcgctctgg 1266
QY 1279 agatattggatattacgacgaagatggaaaatttctttatagttgacgattgaaagaact 1338
Db 1267 tgatatgtcttattatgacaatgatggccattttatatttggtgacaggtgaagtcatt 1326
QY 1339 tattaaatacaaggatatactcaggttgccctgctgaactggaaaatctgctttttacaaca 1398
Db 1327 aattaaatataaagttatcaggttgccactgctgaattgggggaatactcttacaaca 1386
QY 1399 tccaagtattgctgctgggtgttactggagttccggacgaatttggtagacaattacc 1458
Db 1387 tccgtatatgtgtgacgcggttactgggtatataccggatgaagccggtggagcttcc 1446
QY 1459 tgctgctgtgtgtgttagaatctggcaagacgctgactgaaaagggaagttcaagattt 1518
Db 1447 agctgcaggtgtgtgtacagactggaaaatatctaaacgaacaatcgtacaaaattt 1506
QY 1519 tattgcagcacaaagtcactccaacaaaagcatcttcgagcggtgtcgattttgtagacag 1578
Db 1507 tgtttccagtcgaagtttcaacagccaaatggctacgtggtgggtgaaatttttggatga 1566
QY 1579 tattccgaagccctactgaaaaactcatcagaagagctccgagaaaattttg 1634
Db 1567 aattccaaaaggatcaactggaaaaattgacagaaaagtgttaagacaaaatgtttg 1622

Search completed: September 7, 2002, 22:32:12
Job time: 9381 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:04:01 ; Search time 13.45 seconds
(without alignments)
1568.933 Million cell updates/sec

Title: US-09-993-874-2
Perfect score: 2812
Sequence: 1 MEENIRHGERPRDIYVPGS.....KLIRKELREIFAORAPKSKL 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 36719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1584.5	56.3	550	1 LUCI_PHOPY	P08659 photinus py
2	1491.5	53.0	548	1 LUCI_LUCCR	P13129 luciola cru
3	1458.5	51.9	548	1 LUCI_LUCIA	O01158 luciola lat
4	862	30.7	547	1 4CL1_TOBAC	O24145 nicotiana t
5	855	30.4	542	1 4CL1_TOBAC	O24146 nicotiana t
6	847	30.1	545	1 4CL1_SOLTU	P31684 solanum tub
7	838	29.8	545	1 4CL1_SOLTU	P31685 solanum tub
8	836.5	29.7	553	1 4CL1_VANPL	O24540 vanilla pla
9	833.5	29.6	563	1 4CL1_ORYSA	P17814 oryza sativ
10	824	29.3	562	1 4CL2_SOYBN	P31687 glycine max
11	820.5	29.2	556	1 4CL2_ARATH	O98725 arabidopsis
12	810	28.8	544	1 4CL2_PETCR	P14913 petroselinu
13	807.5	28.7	544	1 4CL1_PETCR	P14912 petroselinu
14	807	28.7	561	1 4CL1_ARATH	O42524 arabidopsis
15	773	27.5	537	1 4CL1_PINTA	P41636 pinus taeda
16	763.5	27.2	561	1 4CL3_ARATH	O98777 arabidopsis
17	719.5	25.6	569	1 4CL2_ORYSA	O42982 oryza sativ
18	638.5	22.7	293	1 4CL1_SOYBN	P31686 glycine max
19	615.5	21.9	560	1 LCFA_BACSU	P94547 bacillus su
20	535.5	19.0	561	1 LCFA_ECOLI	P29212 escherichia
21	444.5	15.8	562	1 LCFA_HAEIN	P46450 haemophilus
22	432	15.4	486	1 MENE_BACSU	P23971 bacillus su
23	420	14.9	469	1 MENE_LISMO	P58730 listeria mo
24	414	14.7	543	1 FAT2_YEAST	P38137 saccharomyc
25	406.5	14.5	548	1 YDID_ECOLI	P38135 escherichia
26	406	14.4	469	1 MENE_LISIN	O92548 listeria in
27	404	14.4	522	1 CAIC_ECOLI	P31552 escherichia
28	380.5	13.5	572	1 ACFA_BACSU	P39062 bacillus su
29	378	13.4	546	1 ALRK_PSEOL	O00594 pseudomonas
30	372	13.2	699	1 LCFA_RAT	P18163 rattus norv
31	369.5	13.1	536	1 ENTE_ECOLI	P10378 escherichia
32	365	13.0	492	1 MENE_STAAU	O53634 staphylococ
33	350	12.4	699	1 LCFA_MOUSE	P41216 mus musculu

34	349	12.4	1088	1	TYCA_BACBR	P09095 bacillus br
35	344	12.2	675	1	ACSA_CANAL	O94049 candida alb
36	342	12.2	698	1	LCFB_HUMAN	P33121 homo sapien
37	335.5	11.9	699	1	LCFA_HUMAN	P41215 homo sapien
38	330.5	11.8	6486	1	TYCC_BACBR	O30409 b tyrocidin
39	329.5	11.7	6359	1	BACC_BACLI	O68008 b bacitraci
40	328	11.7	539	1	DHBE_BACSU	P40871 bacillus su
41	323	11.5	698	1	LCFA_CAVPO	O91d6 cavia porce
42	321.5	11.4	683	1	ACSI_KLULA	O60011 kluyveromyc
43	321	11.4	1098	1	GRSA_BACBR	P14687 bacillus br
44	320	11.4	744	1	LCF2_YEAST	P39518 saccharomyc
45	319	11.3	683	1	LCFE_RAT	O88813 rattus norv

ALIGNMENTS

RESULT	1
LUCI_PHOPY	
ID	LUCI_PHOPY
AC	P08659; STANDARD; PRT; 550 AA.
DT	01-JAN-1988 (Rel. 06, Created)
DT	01-JAN-1988 (Rel. 06, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).
OS	Photinus pyralis (North American firefly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC	Elateriformia; Cantharoidea; Lampyridae; Photinus.
OX	NCBI_TaxID=7054;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=87144243; PubMed=38212727;
RA	de Wet J.R., Wood K.V., Deluca M., Helinski D.R., Subramani S.;
RT	"Firefly luciferase gene: structure and expression in mammalian
RL	cells.";
RL	Mol. Cell. Biol. 7:725-737(1987).
RN	[2]
RP	SUBCELLULAR LOCATION.
RX	MEDLINE=87204117; PubMed=3554235;
RA	Keller G.-A., Gould S., de Luca M., Subramani S.;
RT	"Firefly luciferase is targeted to peroxisomes in mammalian cells.";
RN	Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).
RN	[3]
RX	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RP	MEDLINE=96398615; PubMed=8805533;
RX	Conti E., Franks N.P., Brick P.;
RT	"Crystal structure of firefly luciferase throws light on a
RL	superfamily of adenylate-forming enzymes.";
RL	Structure 4:287-298(1996).
RN	[4]
RX	X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RP	MEDLINE=99007339; PubMed=9788915;
RA	Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;
RT	"Structural basis for the inhibition of firefly luciferase by a
RL	general anesthetic.";
RL	Biophys. J. 75:2205-2211(1998).
CC	- - FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.
CC	- - CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP -> oxidized
CC	Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.
CC	- - COPACTOR: MAGNESIUM.
CC	- - SUBCELLULAR LOCATION: Peroxisomal.
CC	- - SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC	FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----

```
DR EMBL: M15077; AAA29795.1; -
DR EMBL: X84848; AAA59283.1; -
DR EMBL: U03687; AAA03561.1; -
DR EMBL: U89934; AAB64396.1; -
DR EMBL: U89935; AAB64399.1; -
DR PIR: A26772; A26772.
DR PDB: ILIC1; 26-MAR-97.
DR PDB: 1BA3; 11-NOV-98.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
FT Peroxisome; 3D-structure.
KW SITE 548 550 MICROBODY TARGETING SIGNAL.
SQ SEQUENCE 550 AA; 60745 MW; E380FCE9D56ACCDE CRC64;

Query Match 56.3%; Score 1584.5; DB 1; Length 550;
Best Local Similarity 55.5%; Pred. No. 4.5e-107;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGAGQOLYQSLYKFAFPE--AIIDAHTEVISAQIFETS 57
DB 1 MEDAKNIKKGPAPPFLEDGTAGEQLHKAMKRYALVPCTIAFTDAHIEVNTIAYEFEMS 60
QY 58 CRLAVSIEOYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSDMYTDCGLTGHINI 117
DB 61 VRLAEMKRYGLNTHRIWVCSENSLQFMPVLGALFTGVAVAPANDIYNRELLNSMI 120
QY 118 SKPTIMFSSKKALPLIRVQNLSPKVVVIDSMDYINGECVSTFVARYTDHTFDPLS 177
DB 121 SQTVPVSVSKKGLQILNVQKKLPIQKIIIMDSKTDYOGFQSGMVTFTVSHLPPGFNEVD 180
QY 178 FPPKDFDPLEKALIMSSSGTGLPKGVLSHRSLTIRFVHSRDIYCTRTVPQTSILSL 237
DB 181 FVPESFDRDKTALIMNSSGTGLPKGVLPHTACTVRSFHSARDPIFGNQIIPDTAILSV 240
QY 238 VPFHAFGMFTLSYFVVGKLVYMLKKEGALFLKTIONYKIPTIWAAPPVNVFLAKSPL 297
DB 241 VPFHGFGMFTLGLYICGFRVVLMYRFEELFLSLQYKIQSALLVPTLFSFAKSTL 300
QY 298 VDQYDLSSLETVATGAPLGKDVAEAVAKRLKPLIGIIGYGLTETCCAVMITPHNAVKTG 357
DB 301 IDKYDLSNLHEIASGAPLSKEVGEAVAKRPHLPGLRGYGLTETTSAILITPEGDKPG 360
QY 358 STGRPLPYKAKVLNATGKALGPGEICQFQSEMIKGYNNPEATIDTDKDGWLHS 417
DB 361 AVGVVVPFPEAKVVDLTGKTLGVNQRGELCYRGPIMSGVYNNPEATNALIDKDGWLHS 420
QY 418 GDIGYDDEGNFFIYDLRLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDFEGGOL 477
DB 421 GDIAFWDEDEHFFIYDLRLKSLIKYGYOVAPAELESILLOHPNIFDAGVAGLPDDAGEL 480
QY 478 PAACVWLESGKTLTEKEVQDTAAQVPTPKHLRGVGVFVDSIPKGTGLKRLKELREIP- 536
DB 481 PAAVVVLEHGKMTKEKIVDVASQVTTAKKLRGVGVFVDEVPKGLTGCKLDARKIREILI 540
QY 537 -AQRAPSKSL 545
DB 541 KAKKGKSKSL 550

RESULT 2
ID LUCI_LUCCR
AC P13129;
DT 01-JAN-1990 (Rel. 13; Created)
DT 01-JAN-1990 (Rel. 13; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

```
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=7051;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2473944;
RX MEDLINE=89326143;
RA Masuda T., Tatsumi H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, Luciola cruciata.";
RL Gene 77:265-270(1989);
CC -!- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -!- CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP -> oxidized Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.
CC -!- COFACTOR: MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M26194; AAA29135.1; -
DR PIR: JS0181; JS0181.
DR HSP: P08659; IBA3.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
KW Peroxisome.
FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60017 MW; 2052D6189E79109F CRC64;

Query Match 53.0%; Score 1491.5; DB 1; Length 548;
Best Local Similarity 53.1%; Pred. No. 2.4e-100;
Matches 289; Conservative 98; Mismatches 154; Indels 3; Gaps 2;

QY 3 ENIRHGERPRDIVHPGAGQOLYQSLYKFAFPE-AIIDAHTEVISAQIFETSCRIA 61
DB 7 DENIVVGPKFPPIEGSGAGTQLRKYMERYAKGAIAETNAVTVGYDYSVAEYLESCGLG 66
QY 62 VSIQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSDMYTDCGLTGHLSKPT 121
DB 67 KALQNYGLVVDGRIALCSENCERTFIPVIAGLFVGVAFTNEIYTLRELHSLGTSKPT 126
QY 122 IMFSSKKALPLILRQOOLNLSFIKKVVVIDSMDYINGECVSTFVARYTDHTFDPLSFTPK 181
DB 127 IVFSSKKGLDKVITVQKTTIKTIVILDSKVDYRGYQCLDTFIKRNTPPGQASFKTV 186
QY 182 FDPDLEKALIMSSSGTGLPKGVLSHRSLTIRFVHSRDIYCTRTVPQTSILSLVPPH 241
DB 187 EYDRKEQVALIMNSSGTGLPKGVLTHTNTVTRSHARDPIYGVQSPCTAVLIVPPH 246
QY 242 HAFGMFTLSYFVVGKLVVYMLKKEGALFLKTIONYKIPTIIVVAPPVNVFLAKSPLVDQY 301
DB 247 HGFGMFTLGLYLCGFRVYMLTKFDEETFLKTLQDYKCTSVILVPTLFAILNKSELLNKY 306
QY 302 DLSSLETVATGAPLGKDVAEAVAKRLKPLIGIIGYGLTETCCAVMITPHNAVKTGSGTGR 361
DB 307 DLSNLVEIASGGAPLSKEVGEAVARRFNLPGVROGYGLTETTSAILITPEGDKFGASCK 366
QY 362 PLPYIKAKVLDNATGKALGPGEICQFQSEMIKGYNNPEATIDTDKDGWLHSGDIG 421
DB 367 VVPLFAKVIDLDTKSLGPNRRGEVCVKPMLKGYNNPEATKELIDEEGWLHTGDIG 426
QY 422 VYDEGNFFIYDLRLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDFEGGOLPAAC 481
DB 422 VYDEGNFFIYDLRLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDFEGGOLPAAC 481
```

```

Db 427 YYDEEKHEFFIVDRLSKLIKYGQVPPAELESVLLQHPISIFDAGVAGVDPDPVAGELPGAV 486
Qy 482 VVLESGKTLTEKEVODFAAQVPTPKHLRGVGVFVDSIPKPGTGKLIKRELREIFAQAP 541
Db 487 VVLESGKNTKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGIDGRAIREIL--KKP 544
Qy 542 KSKL 545
Db 545 VAKM 548

RESULT 3
LUCI_LUCIA
ID LUCI_LUCIA STANDARD; PRT; 548 AA.
AC Q01158;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidae; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A. PubMed-1610896;
RX MEDLINE-92305054; PubMed-1610896;
RA Tatsumi H., Kajiyama N., Nakano E.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA clone
RL encoding luciferase of a firefly, Luciola lateralis.";
RL Biochim. Biophys. Acta 1131:161-165(1992).
CC -!- FUNCTION: PRODUCES GREEN LIGHT.
CC -!- CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP -> oxidized
CC Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.
CC -!- COFACTOR: MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X66919; CAA47358.1; -
DR PIR; S23437; S23437.
DR HSP; P08659; IBA3.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
KW Peroxisome.
FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60125 MW; AC62F9320BB6D4A6 CRC64;

Query Match 51.9%; Score 1458.5; DB 1; Length 548;
Best Local Similarity 51.8%; Pred. NO. 5.8e-98;
Matches 282; Conservative 101; Mismatches 158; Indels 3; Gaps 2;

Qy 3 RENIRHGRPRDIVHPGAGQOLYSLYKFAFPPE-AIIDAHTNEVISYAOIFETSCRLA 61
Db 7 DENIVGPEFFPIIEGSGAGRLYMDRYAKLGAIAFTNAULTGYDYTYAEYLEKSCCLG 66
Qy 62 VSIEQYGLNENNVGCSNNINFFNPVLAALYLGIPVATSNMDYTDGELTGLHNTSKPT 121
Db 67 EALKNYGLVWDCRIALCSENCCEFTIPVLAGLFIGVGAFTNEITLRELHSLGLSKPT 126
Qy 122 IMFSSKKALPLILRVQOONLSFIKKVWVDSMYDINGVECVSTFVARYTDHTFDPLSFTPK 181

```

```

Db 127 IVFSSKKGLDKVITVQKTVTAIKTIVILDSKVDYRGYQSMDFIRKNTPOGFGSGSKTV 186
Qy 182 OFDPILEKIALINSSGTTGLPKGVVLSHRSRLTIRVHGRSDPTYGTRTVPTQTSILSVPPH 241
Db 187 EVNRKEQVALINSSGTTGLPKGVLTENAVTRFESHARDPIYGNVSPGTALLTVPPH 246
Qy 242 HAFGMFTTSLFVVGKLVVVKLAKKFFGALFLKTIQNIKPTIVVAPPVWVFLAKSLVDQY 301
Db 247 HGFNMFTTGLYLTGFCFIRVIMLTKFDEETFLQDYKSSVILVPTFLFALNRSELLDKY 306
Qy 302 DLSLSTEATGAPLGGKVAEAVAKRLKLPGLIIOGVLGTETCCAVMITPHNAVKTGSTGR 361
Db 307 DLSNLVEIASGGAPLSKEIGEAVARFNLPGVQGTGLTETTSALIIITPEGDDKPCASGR 366
Qy 362 PLPYIKAKVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKQWMLHSGDIG 421
Db 367 VVPLFKAKVIDLTKTLGPNRRGVCVKGPMLMGVYDNPENATREIIDEQWMLTGDIG 426
Qy 422 YYDEGNFTYVDRKELIKYGYQVAPAELENLLQHPISADAGVGVDPDEGGQLPAAC 481
Db 427 YYDEEKHEFFIVDRLSKLIKYGQVPPAELESVLLQHPNIFDAGVAGVDPDPVAGELPGAV 486
Qy 482 VVLESGKTLTEKEVODFAAQVPTPKHLRGVGVFVDSIPKPGTGKLIKRELREIFAQAP 541
Db 487 VVLESGKNTKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGIDGRAIREIL--KKP 544
Qy 542 KSKL 545
Db 545 VAKM 548

RESULT 4
4CL1_TOBAC
ID 4CL1_TOBAC STANDARD; PRT; 547 AA.
AC 024145;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate--CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA
DE synthase 1).
GN 4CL1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9641641; PubMed-8819324;
RA Lee D., Douglas C.J.;
RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
RT (4CL) gene family, cDNA structure, gene inheritance and expression,
RT and properties of recombinant proteins.";
RL Plant Physiol. 112:193-205(1996).
CC -!- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -!- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U50845; AAB18637.1; -
DR HSP; P08659; ILCI.
DR Mendel; 24477; Nicotiana1179;24477.

```



```

DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-comarate-CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-comaroyl-CoA
GN synthase 1).
GN 4CL1 OR 4CL-1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridaceae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=91217100; PubMed=2022667;
RX Becker-Andre M., Schulze-Lefert P., Hahlbrock K.;
RT "Structural comparison, modes of expression, and putative cis-acting
RT elements of the two 4-comarate; CoA ligase genes in potato.";
RL J. Biol. Chem. 266:8551-8559(1991).
CC -|- CATALYTIC ACTIVITY: ATP + 4-comarate + CoA = AMP + diphosphate +
CC 4-comaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M62755; AAA33842.1; -.
DR PIR: A39827; A39827.
DR HSP: P08659; JBA3.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
DR DR:
DR Ligase; phenylpropanoid metabolism; Multiogene family.
KW SEQUENCE 545 AA; 59619 MW; DE183683B774BA71 CRC64;
SQ

```

Oy	419	DIGYDEGNPFIVDRUKELIKYGYQVAPAELENLLQHPSTADAGVTGVDFGQLP	478
		: : : : : : : : :	
Db	423	DIGIFIDDDLELFIVDRUKELIKYGFQVAPAELEALLINHPDTSDAVVPMIDEQAGEYP	482
		: : : : : : : : :	
Oy	479	AACVVLSEGGTKLTKVEVDIFAQAQTPTKHLRGVGVFVDSIPKGPTGKLIRKLR	533
		: : : : : : : : : : :	
Db	483	VAFVVRNSNGSTIPEDEVKDFISKOVIFYKKIK-RVFVEVTPKSPSKILRKDLR	536
		: : : : : : : : : : :	
RESULT	7		
ID	4CL2 SOLTU		
ID	4CL2_SOLTU	STANDARD;	PRT; 545 AA.
AC	P31685;		
DT	01-JUL-1993 (Rel. 26, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	4-coumarate--CoA ligase 2 (EC 6.2.1.12) (4CL 2) (4-coumaroyl-CoA synthase 2).		
DE	NCBI Taxid=4113;		
GN	4CL2 OR 4CL-2.		
OS	Solanum tuberosum (Potato).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
OC	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.		
OX	NCBI_Taxid=4113;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91217100; PubMed=2022667;		
RA	Becker-Andre M., Schulze-Lefert P., Hahlbrock K.;		
RT	"Structural comparison, modes of expression, and putative cis-acting elements of the two 4-coumarate; CoA ligase genes in potato.";		
ET	J. Biol. Chem. 266:8551-8559(1991).		
RL	-1- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate + 4-coumaroyl-CoA.		
CC	-1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.		
CC	-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.		
CC			
PIR	B39827; B39827.		
DR	HSSP; P08659; 1BA3		
DR	InterPro; IPRO00873; AMP-bind.		
DR	Pfam; PF00501; AMP-binding; 1.		
DR	PROSITE; PS00455; AMP_BINDING; 1.		
DR	Ligase; Phenylpropanoid metabolism; Multigene family.		
DR	SEQUENCE 545 AA; 53625 MW; 5481F0B0AFEAF39E0 CRC64;		
SW	KW		

Db 306 RTVMGSAAPLGKLEDAV--RAKFNNAKLGQGYGTEAGPVAMCLAFAPKEPFD-IKSGA 362

QY 359 TGRPLPYIAKAVLDNATKALGPGERGEICFQSEIMKGYNNPEATIDTIDKQGLHSG 418

Db 363 CQTVVNAEMKIVDPDTCGLPRNQPGETCIRGQIMKGYLNDPEATARTIEKGLHTG 422

QY 419 DIGYVDEGNFFIVDLRLKELIKYQVAPAELEMLLQHPISADAGVGVDFEGGOLP 478

Db 423 DIGFIDDDDELFIIVDLRLKELIKYQVAPAELEALLNHDPISDAAVVPMIDEQAGEVP 482

QY 479 ACNVLESCKTITEKEVQDFIAAQTPTPKHLRGVVFVDSIPKGTGKILRELK 533

Db 483 VAFVYRSNGSTITEDEVKDFISKQVIFYKRIK-RVFFVETVPKSPGKILRKDLR 536

RESULT 8

4CL_VANPL STANDARD; PRT; 553 AA.

AC 024540;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 4-coumarate--CoA ligase (EC 6.2.1.12) (4CL) (4-coumaroyl-CoA synthase).

GN 4CL.

OS *Vanilla planifolia* (Vanilla).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;

OC Vanilloid clade; Vanillinae; Vanilla.

OX NCBI_TaxID=51239;

RN [1]

RP SEQUENCE FROM N.A.

RA Brodelius P., Xue Z.T.;

RT *Isolation and characterization of a cDNA from cell suspension

RT cultures of *Vanilla planifolia* encoding 4-coumarate: coenzyme A

RT ligase.*;

CC Plant Physiol. Biochem. 35:497-506(1997).

CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate + 4-coumaroyl-CoA.

CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID

CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.

CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X75542; CAA53230.1; -.

DR HSSP; P08659; ILCI.

DR Mendel; 24370; Vanpl; 1179; 24370.

DR InterPro; IPR000873; AMP-bind.

DR Pfam; PF00501; AMP-binding; 1.

DR PROSITE; PS00455; AMP-BINDING; 1.

KW Ligase; Phenylpropanoid metabolism.

SQ SEQUENCE 553 AA; 60095 MW; 9A2D250BC84EALCB CRC64;

Query Match 29.7%; Score 836.5; DB 1; Length 553;

Best Local Similarity 35.9%; Pred. No. 5.4e-53;

Matches 198; Conservative 102; Mismatches 177; Indels 75; Gaps 14;

QY 26 YQSLYKFAFPBAIIDAHNTNEISYAQIFETWSCRVAISYEQVGLNENNVGVCSNNINF 85

Db 36 FENISKFSRP-CLINGATDEFTVADVBLISRRVSGSLKLGKIQGDTIMLLPNSPEF 94

QY 86 FNPVLAALYLGIPVATSDNMYTGDGLTGHNLNISKPTIFESSKKALPLILRVQNLSPFIK 145

Db 95 VFAFLGASFIGSISTWNPFPFTSTVIKQAKASNAKLI-----ITQ 135

QY 146 VVVIDSMYDI---NGVECVSTFVARYTDTTDFDPLSF---TPKDFDPLEKIAL-----I 192

Db 136 GCVYDKVKDYACENGKVIISIDTTTTADDAANILHFSELTSAGADENPKVEISPDGVVAL 195

QY 193 MSSSTTTGLPKGVVLSHRSLTIR-----FVHSRDPYGYTRVTPQTSILSLVPF 240

Db 196 PYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYMHSDD-----VLLCVLPL 244

QY 241 HHAFGMFTTSLVFFVGLK---VVMKKPEGALFLTKTQNYKIPTIVVAPPVVMVFLAKSP 296

Db 245 PHYSLSNVL---LCGLRAGSGILIMQKPEIVPPELQKRYKVTIGPFPVPIVLAIAKST 301

QY 297 LVQDYDLSLSEVATGGAPLGRKDAEAVAKRLKLPG--IIQGYGLTEN-----CCAVMI 348

Db 302 VVDNYDLSLVRTVMGSAAPLGKLEDAV--RAKFNNAKLGQGYGTEAGPVAMCLAFAPK 359

QY 349 TPHNAVKTSGTGRPLPYIAKAVLDNATKALGPGERGEICFQSEIMKGYNNPEATIDT 408

Db 360 EPFD-IKSGAGTVMVNAEMKIVDPETGSSLPNRNHPGEICIRGQIMKGYLNDPEATART 418

QY 409 IDKQGLHSGDIGYVDEGNFFIVDLRLKELIKYQVAPAELEMLLQHPISADAGVGTG 468

Db 419 IDKQGLHTGDIGYIDDDDELFIIVDLRLKELIKYQVAPAELEALLTHPCISDAAVVP 478

QY 469 VDFEGGOLPAACVVLVLESCKTITEKEVQDFIAAQTPTPKHLRGVVFVDSIPKGTGKLI 528

Db 479 MKDEAAGEVPVAFVYRSNGSTITEDEVKDFISKQVIFYKRI-NRVFFVEATPKAPSGKITL 537

QY 529 RKELREIFAQRA 540

Db 538 RKDLRLARLAAA 549

RESULT 9

4CL_ORYSA STANDARD; PRT; 563 AA.

AC P17814;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE 4-coumarate--CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA synthase 1).

GN 4CL1 OR 4CL.

OS *Oryza sativa* (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV JAPONICA;

RX MEDLINE=91045096; PubMed=2235510;

RA Zhao Y., Kung S.D., Dube S.K.;

RT *Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL.1.*;

RL Nucleic Acids Res. 18:6144-6144(1990).

CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate + 4-coumaroyl-CoA.

CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID

CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.

CC -|- INDUCTION: BY FUNGAL ELICITOR AND UV IRRADIATION.

CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X52623; CAA36850.1; -.

```

DR PIR: JU0311; JU0311.
DR HSP; P08659; ILCI.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Ligase; Phenylpropanoid metabolism; Multigene family.
SQ SEQUENCE 563 AA; 60841 MW; 358329401C6B44B9 CRC64;

Query Match      29.6%; Score 833.5; DB 1; Length 563;
Best Local Similarity 39.3%; Pred. No. 9.1e-53;
Matches 207; Conservative 89; Mismatches 192; Indels 39; Gaps 16;

QY 30 YKFAFPE-----AIDAHTNEVISYAQIFETSCRLAVSIEQ--YGLNENNVVGVSENN 82
DB 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 42 YCFERLPEVAARPCLDGATGVTYADVRLSRRLAARLAPLGLRGVVMSSLLRNS 101
QY 83 INFENPVALALYLGIPVATSDMYTDELTHLNTSKPTIMFSSKALPLILRVOONLSF 142
DB 102 PEFVLSFFAASRVGAATYANPSTPHETESQLAAGATVVTESMAAD--KLPSSHSG 158
QY 143 IKKVVVDSMYDINGVECVSTFYARTDHTFDPFSFPKD---FDPLEKIALIMSSGTT 199
DB 159 ALTVVLIDERRD---GCLHFWDLMSEDEASPLAGDEDEKVFDPDDVVAL-PYSSGTT 213
QY 200 GLPKGVVLSHRSLT---IRFVHSRDIYGTTRVPQTSILSLVPFHAFGCFHTLSYFV-Y 255
DB 214 GLPKGVMLTHRSLSVTSVAQOVDPGENPNGLHA--GDVILCALPMPFIHYSLNTIMCGLRV 271
QY 256 GLKVVMLKFEKALFKLTONTKYPIVAVPMVPLAKSPLVDQVDSLSEVATGGAP 315
DB 272 GAAIVVMRRFDLAAMDLVHRHVITIAPVPIVAVAKSEAAARDSLSSVMVLSGAAP 331
QY 316 LKQDVAEAVAKRLKPLGII--QCYGLTET-----CCAVMITPHNAVKTGSTGRPLPYTK 367
DB 332 MGKDIEDAF--NAKLPGAVLGOGYGTAGPVLSCMLAEAKPEFK-VKSGACGTVVVRNAE 388
QY 368 AKVLNDATCKALGPGERGEICFOSEIMKGYNNPEATIDTDKDWLHSGDIGYDEOG 427
DB 389 LKIIDPTCKSLGRNURGEICIRGOIMKGYLNNPEATKNTIDAEGLWHTDGIVYDDDD 448
QY 428 NFFVDRKLKIKYQYVAPAELENLLQHPISADAGVTGPDFFGGOLPAACVVLESG 487
DB 449 ELFIVDRLEIKYRFQVAPAELEALLNTHPSIADAANVGL--KP-GEIPVAFVAKTEG 505
QY 488 KTLTEKEVQDFAAQVTPTKHLRGVGVFVDSIPKGPSTGKILRKELRE 534
DB 506 SELSEDDVQFVAKEVIYKKIR-EVFFVDKIPKAPSGKILRKELRK 551

RESULT 10
4CL2_SOYBN STANDARD; PRT; 562 AA.
AC P31687;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate-CoA ligase 2 (EC 6.2.1.12) (4CL 2) (4-coumaroyl-CoA
DE synthase 2) (Clone 4CL16).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_Taxid=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Lindermayr C.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 140-562 FROM N.A.
RC STRAIN=CV. HARISOY 63;
RX MEDLINE=94105342; PubMed=8278545;
RA Uhlmann A., Ebel J.;

```

```

RT "Molecular cloning and expression of 4-coumarate:coenzyme A ligase,
RT L.) against pathogen attack.";
RL Plant Physiol. 102:1147-1156(1993).
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69955; CAC36095.1; -.
CC PIR; S31709; S31709.
CC PIR; PQ0772; PQ0772.
CC HSP; P08659; ILCI.
CC InterPro: IPR000873; AMP-bind.
CC Pfam: PF00501; AMP-binding; 1.
CC PROSITE: PS00455; AMP_BINDING; 1.
KW Ligase; Phenylpropanoid metabolism; Multigene family.
SQ SEQUENCE 562 AA; 60963 MW; 2AB4652DDA5160B2 CRC64;

Query Match      29.3%; Score 824; DB 1; Length 562;
Best Local Similarity 36.5%; Pred. No. 4.4e-52;
Matches 198; Conservative 100; Mismatches 192; Indels 52; Gaps 13;

QY 26 YOSLYKFAFPEAIDAHTNEVISYAQIFETSCRLAVSIEQYGLNENNVGVSENNIF 85
DB 47 FQNLQSFARHP-CLIVGPASKTFYADTHLSSKIAAGLSNLGILKGDVVMILLQNSADF 105
QY 86 FNPVLAALYLGPVATSDMYTDELTHLNTSKPTIMFSSKALPLILRVOONLSFIKK 145
DB 106 VFSFLAISMIGAVATTANPFYTAPEIFKQFTVSAKALI-----ITQ 146
QY 146 VVVIDSMYDINGVECVSTF-VARYTDHTDPLSFT-----PK-DFDPLEKIALIM 193
DB 147 AMYVDKLRNHDGAKLGEDFKVTVDDPPENCNLFHVSLSSEANESDVPEVEIHPDDAVMPF 206
QY 194 SSGTGTGLPKGVVLSHRSLTIRFVHSRD----PIYGTTRVTPQTSILSLVPFHAFGFTT 249
DB 207 -SSGTTGLPKGVILTHKSLTTSVAQOVDCENPNLYLT---TEDVLLCVLPFLHFISLSV 262
QY 250 LSYFV-VGLKVVMLKKFEGALFKLTQNTKYPIVAVPMVPLAKSPLVDQVDSLSTE 308
DB 263 LLCALRAGSAVLLMQKFEIGTLLLELQIRHRVSVAMVVPPLVLALAKNPMVADFDLSIRL 322
QY 309 VATGGAPLCKDVAEAVAKRLKPLGIIQGYGLTET-----CCAVMITPHNAVKTGSGRP 362
DB 323 VLSGAAPLKEDEEALRNRMPOAVLGQGYGTEAGPVLSCMLGFAKQFPQ-THSGSGTV 381
QY 363 LPYIKARVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTDKDWLHSGDIGY 422
DB 382 VRNAELKVVDPEPGRSLGYNQPGCEICIRGOQINKGYLNDAAATASTIDSEGLWHTGDVG 441
QY 423 YDEGNGFFIVDRKELIKYQYVAPAELENLLQHPISADAGVTGPDFFGGOLPAACV 482
DB 442 VDDDDIEIFVDRVKELIKYQYVAPAELEGLVSHPSIADAANVVPKQVAAAGEVPAFV 501
QY 483 VLESAGTLTEKEVQDFAAQVTPTKHLRGVGVFVDSIPKGPSTGKILRKELR---EIFAOR 539
DB 502 VRSNGFDLTEEAVKETAKQVVFYKRLH-KYFVFAIHPKSPSGKILRKELRAKLEATNQ 560
QY 540 AP 541
DB 561 TP 562

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----

DR EMBL; X13325; CAA31697.1; -;
DR EMBL; X05351; CAA28960.1; -;
DR PIR; S01667; S01667;
DR PIR; S15695; S15695;
DR HSSP; P08659; ILCI.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Phenylpropanoid metabolism; Multigene family.
SQ SEQUENCE 544 AA; 59783 MW; B477965C68F8C534 CRC64;

Query Match 28.8%; Score 810; DB 1; Length 544;
Best Local Similarity 34.8%; Pred. No. 4.3e-51;
Matches 189; Conservative 105; Mismatches 167; Indels 82; Gaps 13;
QY 39 IDAHTNEVISVAQIFETSCRLAVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGTIP 98
DB 45 LINGATGETFTYSQVELLSRKVASGLNKLGIQOGDTIMLLPNSPEYFFAFGLGAYRGAI 104
QY 99 VATSNDMTDGLTGHNLISKPTIMFSSKKALPLILRVQONLSFIKKVVVIDSMYDINGV 158
DB 105 STMANPFTSAEVIKQLKAS-----LAKLIITQACY---V 136
QY 159 ECVSTFVAR-----YDHTFDPLSFT-----PKDFDPLEKIALINSSSGTTGL 201
DB 137 DKVKDYAAEKNTIICIDDAPODCLRFKLMDEASEMPEVVVDSDDVVALPYSSGTTGL 196
QY 202 PKGVVLSHRSLTIR-----FVHSRDPITYGTRTPOTSILSLVPFHFAFGMFTT 249
DB 197 PKGVMLTHKGLVTSVAQQVGDGNPNLYMHSD-----VMICILPLFIHYSLNAY 245
QY 250 LSYFV-VGLKVVMLKKFEGALFKTIQNYKIPTIVVAPPVVMVFLAKSPLDQYDLSLSTE 308
DB 246 LCCGLRAGVTILIMQKFDIVPELEIQRKVTIGFPVPIVLAIAKSPVVDKYDLSVVRT 305
QY 309 VATGAPGLKQVAEAVAKRLKLPG--IIQGYLGTET-----CCAVMTIPTHNAVKTGSTG 360
DB 306 VMSGAPLGLKELEDV--RAKFPNKLGGQYGMTEAGPVLAMCLAPAKEPYE-IRSGACG 362
QY 361 RPLPIKAVLDNATGKALGPCERGEICFQSEIMKMGYNPNPEATIDTIDKDWLHSGDI 420
DB 363 TVVRNAEMKIVDPNATSLPRNGRCEICRGQIMKGYLNDPESTRTTIDEEGLHTGDI 422
QY 421 GYDDEGFFIVDLRELKIKYGYVAPAELENLLIQLHPSTADAGVTGVPDFEGQLPAA 480
DB 423 GFIDDDDELIVDLRELKIKYGFQVAPAELEALLTHTPTISDAAVVPMIDEKAGEVPA 482
QY 481 CVVLSEGTLEKEVODFTAAQVTPKHLRGVGVFVDSIPKPGTQKLRKELR-EIFAQR 539
DB 483 FVVRTNGFTTBEELIKQFVSKOVVFKRI-FRVFFVDAIPKSPGKILRKDLRAKIASGD 541
QY 540 APK 542
DB 542 LPK 544

RESULT 13

ID 4CL1_PETCR STANDARD; PRT; 544 AA.
AC P14912;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 4-coumarate-CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA
synthase 1).
GN 4CL1 OR 4CL-1.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

CC Asteridae: euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID=4043;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005119; PubMed=3169018;
RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
RA Hahlbrock K.;
RT *Primary structures and catalytic properties of isoenzymes encoded by
RT the two 4-coumarate:CoA ligase genes in parsley.*;
RL Eur. J. Biochem. 176:661-667(1988).
RN [2]
RP SEQUENCE OF 1-8 FROM N.A.
RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
RT *Structure and elicitor or U.V.-light-stimulated expression of two
RT 4-coumarate:CoA ligase genes in parsley.*;
RL EMBL J. 6:1189-1195(1987).
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- INDUCTION: BY FUNGAL ELICITOR AND UV IRRADIATION.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; X13324; CAA31696.1; -;
DR EMBL; X05350; CAA28959.1; -;
DR PIR; S01667; S01667.
DR HSSP; P08659; IBA3.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Phenylpropanoid metabolism; Multigene family.
SQ SEQUENCE 544 AA; 59825 MW; 22BBAD78F25D0C8 CRC64;

Query Match 28.7%; Score 807.5; DB 1; Length 544;
Best Local Similarity 34.9%; Pred. No. 6.6e-51;
Matches 186; Conservative 103; Mismatches 163; Indels 81; Gaps 12;
QY 39 IDAHTNEVISVAQIFETSCRLAVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGTIP 98
DB 45 LINGATGETFTYSQVELLSRKVASGLNKLGIQOGDTIMLLPNSPEYFFAFGLGAYRGAI 104
QY 99 VATSNDMTDGLTGHNLISKPTIMFSSKKALPLILRVQONLSFIKKVVVIDSMYDINGV 158
DB 105 STMANPFTSAEVIKQLKASQ-----AKLIITQACY---V 136
QY 159 ECVSTFVAR-----YDHTFDPLSFT-----PKDFDPLEKIALINSSSGTTGL 201
DB 137 DKVKDYAAEKNTIICIDDAPODCLRFKLMDEASEMPEVVVDSDDVVALPYSSGTTGL 196
QY 202 PKGVVLSHRSLTIR-----FVHSRDPITYGTRTPOTSILSLVPFHFAFGMFTT 249
DB 197 PKGVMLTHKGLVTSVAQQVGDGNPNLYMHSD-----VMICILPLFIHYSLNAY 245
QY 250 LSYFV-VGLKVVMLKKFEGALFKTIQNYKIPTIVVAPPVVMVFLAKSPLDQYDLSLSTE 308
DB 246 LCCGLRAGVTILIMQKFDIVPELEIQRKVTIGFPVPIVLAIAKSPVVDKYDLSVVRT 305
QY 309 VATGAPGLKQVAEAVAKRLKLPG--IIQGYLGTET-----CCAVMTIPTHNAVKTGSTG 360
DB 306 VMSGAPLGLKELEDV--RAKFPNKLGGQYGMTEAGPVLAMCLAPAKEPYE-IRSGACG 362
QY 361 RPLPIKAVLDNATGKALGPCERGEICFQSEIMKMGYNPNPEATIDTIDKDWLHSGDI 420

This Page Blank (uspio,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: September 4, 2002, 15:02:01 ; Search time 20.93 Seconds
(without alignments)
2502.087 Million cell updates/sec

Title: US-09-993-874-2
Perfect score: 2812
Sequence: 1 MEENIRHGERPRDIHVPGS.....KLIRKELREIFAQAPKSKL 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1584.5	56.3	550	1 A26772	Photinus-luciferin
2	1534	54.6	547	2 S62787	Photinus-luciferin
3	1491.5	53.0	548	1 J50181	Photinus-luciferin
4	1483	52.7	548	2 S33788	Photinus-luciferin
5	1458.5	51.9	548	1 S23437	Photinus-luciferin
6	1393	49.5	543	2 S29354	Photinus-luciferin
7	1388	49.4	543	2 S29355	Photinus-luciferin
8	1382	49.1	543	2 S29352	Photinus-luciferin
9	1377	49.0	543	2 S29353	Photinus-luciferin
10	864.5	30.7	542	1 T02074	4-coumarate--CoA 1
11	855	30.4	542	2 T03789	4-coumarate--CoA 1
12	847	30.1	545	1 A39827	4-coumarate--CoA 1
13	847	30.1	548	2 T07908	4-coumarate--CoA 1
14	838	29.8	545	1 B39827	4-coumarate--CoA 1
15	833.5	29.6	563	1 J00311	4-coumarate--CoA 1
16	823.5	29.3	557	2 T07909	4-coumarate--CoA 1
17	820.5	29.2	544	2 H85064	4-coumarate--CoA 1
18	815.5	29.0	535	2 T08074	4-coumarate--CoA 1
19	810	28.8	544	2 S15695	4-coumarate--CoA 1
20	807.5	28.7	544	1 S01670	4-coumarate--CoA 1
21	807	28.7	561	2 S57784	4-coumarate--CoA 1
22	781	27.8	537	2 T09755	4-coumarate--CoA 1
23	780	27.7	544	2 T02741	hypothetical prote
24	777.5	27.6	570	2 T08075	4-coumarate--CoA 1
25	777	27.6	1549	2 D86338	protein F5M15.18
26	773	27.5	537	2 T09710	4-coumarate--CoA 1
27	763.5	27.2	561	2 D96674	hypothetical prote
28	752.5	26.8	566	2 F85214	4-coumarate--CoA 1
29	752.5	26.8	566	2 T05038	4-coumarate--CoA 1

30	746.5	26.5	542	2 B96654	hypothetical prote
31	743	26.4	423	2 P00772	4-coumarate--CoA 1
32	719.5	25.6	569	2 T03390	hypothetical prote
33	668.5	23.8	540	2 T18841	hypothetical prote
34	639.5	22.7	566	2 T18607	4-coumarate--CoA 1
35	638.5	22.7	293	2 S31705	4-coumarate--CoA 1
36	628.5	22.4	565	2 D88197	protein ZK1127.2
37	618	22.0	564	2 G84037	long-chain fatty-a
38	615.5	21.9	560	2 D89649	probable long-chain
39	610	21.7	566	2 H84037	long-chain fatty-a
40	606	21.6	593	2 E69378	probable acid--CoA
41	568	20.2	548	2 D69187	probable acid--CoA
42	566.5	20.1	577	2 H69354	probable fatty-ac
43	560.5	19.9	554	2 T25902	hypothetical prote
44	551.5	19.6	638	2 T28932	probable 4-coumar
45	548.5	19.5	569	2 C69471	probable fatty-ac

ALIGNMENTS

RESULT 1
A26772
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - common eastern
N:Alternate names: firefly luciferase
C:Species: Photinus pyralis (common eastern firefly)
C:Date: 19-Nov-1988 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
C:Accession: A26772
R:de Wet, J.R.; Wood, K.V.; DeLuca, M.; Hellinski, D.R.; Subramani, S.
Mol. Cell. Biol. 7, 725-737, 1987
A:Title: Firefly luciferase gene: structure and expression in mammalian cells.
A:Reference number: A26772; MUID:87144243
A:Accession: A26772
A:Molecule type: DNA
A:Residues: 1-550 <DE>
A:Cross-references: GB:M15077; NID:g160793; PIDN:AAA29795.1; PID:g160794
A:Note: the authors translated the codon CAA for residue 134 as Glu
C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,
pecies from green to yellow.
C:Genetics:
A:Introns: 42/3; 112/1; 223/2; 330/2; 448/3; 504/1
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
F:71-535/Domain: acetate--CoA ligase homology <ACL>
F:548-550/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif

Query Match	56.3%	Score	1584.5	DB 1	Length	550			
Best Local Similarity	55.5%	Pred. No.	5.2e-105	Mismatches	137	Indels	5	Gaps	3
Matches	305	Conservative	103						
QY	1	MEE-ENIRHGERPRDIHVPGSAGQQLYQSLYKFASPE--AIIDAHTNEVISYAQIFETS	57						
Db	1	MEDAKNIKGPAFPYLEDGTAGEQLHKAMRYALVPGTIAFTDAHIEVNIYAEVFE	60						
QY	58	CRDAVSIEQYGLNENNVGVCSENNINFPVLAALYGLIPVATSDNDYTDGELTGLN	117						
Db	61	VRLAEAMKRYGLNTHRVVCSSENSLOFPMPVLGALFVAVAPANDIYNERELLSM	120						
QY	118	SKPTIFESSKKALPLLRLVOONLSFIKKVVVIDSMYDINGVCYSTFVARYDHTDPL	177						
Db	121	SOPTVVFVSKGKGLKILNVOKKLPIIQIIMDSKTDYGFQSGMYFTVSHLP	180						
QY	178	FTPKRDPDLKIALIMSSGTTGLPKGVLSHRSLSLIRFVHSRDPYIGTTRTPQOTSIL	237						
Db	181	FVPESFDRDKTIALIMSSGSGTGLPKGVLPHTACVRSFARDPIFGNQLIPDTAIL	240						
QY	238	VFPFHAFGFTLTSYFVVGKVMKFKFEGALFLKFTIQNIKIPTIYVAVPVMVFLAK	297						
Db	241	VFPFHGFGFTTLCYLCIGFRVVMYRFEELFLRSLDQYKIQSALLVPTLFSFFAK	300						
QY	298	VQOYDLSLSTEVATGAPLGKQVAEAVAKRLKPLGIIQGYGLTETCCAVMTPHNA	357						

Db 301 IDKYDLSNLHETASGAPLSKEVGEAVAKRFLHPLGIRGOCYGLTETTSAILITPEGDDKPG 360
QY 358 STGRPLPYIKAKVLNATGKALGPCGERGICQSEMIKNGYNNPEATIDTDKDGWLHS 417
Db 361 AVGVVPPFEAKVVDLTGKTLGVNQRGELCVRGPMISGYNNPEATNALIDKDGWLHS 420
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFEGGL 477
Db 421 GDIAWDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVQDTAAQVTPTKHLRGGVVFVDSIPKPTGKLRKELREIF- 536
Db 481 PAAYVVLHGKTMTEKEIVDVYASQVTTAKLRGGVVFVDEVPKGLTGKLDARKIREILI 540
QY 537 -AORAPKSKL 545
Db 541 KAKCGKSKL 550
RESULT 2
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Lampyris noctiluca
N:Alternate names: firefly-type luciferase
C:Species: Lampyris noctiluca
C:Date: 23-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 26-May-2000
R:Salal-Newby, G.B.; Thomson, C.M.; Campbell, A.K.
Biochem. J. 313, 761-767, 1996
A:Title: Sequence and biochemical similarities between the luciferases of the glow-worm
A:Reference number: S62787; MUID:96190714
A:Accession: S62787
A:Molecule type: mRNA
A:Residues: 1-547 <SAL>
A:Cross-references: EMBL:X89479; NID:g899314; PIDN:CAA61668.1; PID:g899315
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
F:71-535/Domain: acetate--CoA ligase homology <ACL>
Query Match 54.6%; Score 1534; DB 2; Length 547;
Best Local Similarity 53.5%; Pred. No. 2c-101;
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;
QY 1 MEE-ENIRHGERPRDIVHPGSGAQQOYLSLYKFAFPE--AIIDAHTNEVISYAIQIFETS 57
Db 1 MEDAKNIMHGPAPFPYLEDGTAGEQLHKMKRYAQVPGTIAPTDAHAENVITYSEVFEMA 60
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATSDMYTDGELTCHLNI 117
Db 61 CRLAETMKRYGLGLQHHIAVCSENSLOQFMPCGALFTGVGVASTNDIYNERELNSLSI 120
QY 118 SKPTIMFSKKALPLTLRVOONLSIKKVVVTDSDMYDINGVECVSTFVARYTDHTFDPLS 177
Db 121 SQTIVSCSKRALQKILGVQKLPPIQKIVILDSREDYMGKQSMYSFTEHSLPAGFNEYD 180
QY 178 FPPKDFPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRVPTQSILSL 237
Db 181 YIPDSFRETATALIMNSSGSTGLPKGVLTQNVCVRFSHCRDPVFGNQIIPDTAILTV 240
QY 238 VPFHAFGCFHTLSFVVGVLKVMKKFEGALFLKTIQNYKPIPTIVAPPVWVFLAKSPL 297
Db 241 IPFHGFCMFTTGLYTCGFRVLVYRFEELFSLRSLQDYKTSALLVPTLFSFFAKSTL 300
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 VDXYDLSNLHETASGAPLAKVEGCAVAKRFLKPLGIRGOCYGLTETTSAILITPEGDDKPG 360
QY 358 STGRPLPYIKAKVLNATGKALGPCGERGICQSEMIKNGYNNPEATIDTDKDGWLHS 417
Db 361 ACQGVVPPFEAKIVDLDTGKTLGVNQRGELCVKGPIMKNGYNNPEATSALIDKDGWLHS 420
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFEGGL 477
Db 481 PAAYVVLHGKTMTEKEIVDVYASQVTTAKLRGGVVFVDEVPKGLTGKLDARKIREILI 540

Db 421 GDIAYYDKDGHFFIVDRKLSLIKYGQVPPAELESILLQHPFFIFDAGVAGIPDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVQDTAAQVTPTKHLRGGVVFVDSIPKPTGKLRKELREIFA 537
Db 481 PAAYVVLSESGKTMTEKEVQDVIYAGQVTKAKLRGGVVKFVDEVPKGLTGKLDARKIREIL- 539
QY 538 ORAPKSKL 545
Db 540 WMGKSKL 547
RESULT 3
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Genji firefly
N:Alternate names: firefly luciferase
C:Species: Luciola cruciata (Genji firefly)
C:Date: 07-Jun-1990 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
C:Accession: JS0181
R:Masuda, T.; Tatsumi, H.; Nakano, E.
Gene 77, 265-270, 1989
A:Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly,
A:Reference number: JS0181; MUID:89326143
A:Accession: JS0181
A:Molecule type: mRNA
A:Residues: 1-548 <MAS>
A:Cross-references: GB:M26194; NID:g159050; PIDN:AAA29135.1; PID:g159051
C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,
pecies from green to yellow.
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
F:73-537/Domain: acetate--CoA ligase homology <ACL>
Query Match 53.0%; Score 1491.5; DB 1; Length 548;
Best Local Similarity 53.1%; Pred. No. 2.2e-98;
Matches 289; Conservative 98; Mismatches 154; Indels 3; Gaps 2;
QY 3 EENIRHGERPRDIVHPGSGAQQOYLSLYKFAFPE--AIIDAHTNEVISYAIQIFETSCRLA 61
Db 7 DENIVVGPKFPPIEGSGAGTOLRKMYERYAKLGAIAFTNAVTVGDVSYAEVLESCLG 66
QY 62 VSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATSDMYTDGELTCHLNIKPT 121
Db 67 KALQNYGLVVDGRIALCSENCPEFPVPIAGLFIGVGVAPTNEIYTLRELVLHSLGSKPT 126
QY 122 IMFSSKKALPLTLRVOONLSIFIKVVVTDSDMYDINGVECVSTFVARYTDHTFDPLSPTK 181
Db 127 IVFSSKKGLDKVITVQKTVTTIKTIVILDSKVDYKGYOCLDTFIKRNTPPGQASFKTV 186
QY 182 DFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRVPTQSILSLVFPFI 241
Db 187 EVDREKEQVALIMNSSSGTGLPKGVQLTHENTVTRFSHARDPIYGNQVSGTAVLTVPFH 246
QY 242 HAFGFTTSLFVVGVLKVMKKFEGALFLKTIQNYKPIPTIVAPPVWVFLAKSPLVQY 301
Db 247 HGFGMFTTGLYLCIGFRVVMVLTTFDEETFLKTLQDYKCTSVILVPTLFAILNKSELLKY 306
QY 302 DLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTGSTGR 361
Db 307 DLSNLVEIASGAPLAKVEGCAVAKRFLNLPVGRVQGTGTETTSAILITPEGDDKPGASK 366
QY 362 PLPYIKAKVLONATGKALGPCGERGICQSEMIKNGYNNPEATIDTDKDGWLHSGDTG 421
Db 367 VYPLFAKVIDLDTKSLGPNRRGEVCKGPMKMGYNNPEATKELIDEGLHTGDI 426
QY 422 YDEGNEFFIVDRKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFEGGLPAAC 481
Db 427 YDESKHEFFIVDRKLSLIKYGQVPPAELESILLQHPISIFDAGVAGVDPVAGELPGAV 486
QY 482 VYLESKTLTEKEVQDTAAQVTPTKHLRGGVVFVDSIPKPTGKLRKELREIFAQAP 541
Db 487 VYLESKNTEKEVMDYVASQVSNAKRLRGGVRFVDEVPKGLTGKLDARKIREIL--KAP 544

Db 545 VAKM 548

RESULT 6

Protein-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YE [validated] - lum
N:Alternate names: firefly-type luciferase
C:Species: Pyrophorus plagiophthalmus
C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000
C:Accession: S29354

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989
A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence of
A:Reference number: S29352; MUID:89242142

A:Accession: S29354

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 49.5%; Score 1393; DB 2; Length 543;

Best Local Similarity 48.9%; Pred. No. 2.2e-91;

Matches 264; Conservative 113; Mismatches 161; Indels 2; Gaps 1;

QY 3 EENIRHGERPRDIVHPGSAQOOLYQSLYKFASFPEALIDAHTNEVISYQAIFETSCRLAV 62

DB 5 ERNVIYGPPLPLEDKTAGEMLFRAIRKHSHPQAIQVDFGDESLSYKEFFETACLLAO 64

QY 63 STEQYGLNENNVVGCSENNINFPNVAALYLGIPVATSDMYTDGELTGHNLISKPTI 122

DB 65 SLHNCGYKNDVVSICAENKREFPIIAAWYIGMIVAPVNESYIPDELCKYMGISKPOI 124

QY 123 MFSSKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARYTDHTFDPLSTPKD 182

DB 125 VFCTKNILNKVLEVSQRTNFIKRIIILDTVENIHGCESLPNFISRYSDGNI--ANFKPLH 182

QY 183 FQPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPITVYGRTPVTPOTSILSLVPPFH 242

DB 183 YDPVEOVAALICSSGTGLPKGMVQTHQICVRLIHAIIDPEAGTQIIPGVTVLVVYPPFH 242

QY 243 AFGMFTTSLYFVVGKLVMLKFEAGFLTKTQNYKIPTIVVAPPVVMFLAKSPLVDQYD 302

DB 243 AFGESINLGYFVGLRVIMLRREFQEAFLKAIQDYEVRSIVNVVPAIILFLSKSPLVDKYD 302

QY 303 LSLTEVATGGAPLCKDVAEAVAKRLKLPGLIIOGYGLTETCCAVMTTPHNAVKTGSTRP 362

DB 303 LSSRELCCGAAPLAKAEVAEIAVKRLNLPICRGFGLTESTSANIHSLGDEFKSGSLGRV 362

QY 363 LPYIKAKVLNDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTIDKGMHLSGDIGY 422

DB 363 TPLMAAKIADRTGKALGPVQGEICIKGPMVSKGVVNNVATKEAIDDDGMHLSGDFGY 422

QY 423 YDEGNFFITVDRLKELIKYQVAPAELENLLQHPISADAGVTGVPDFGGQLPAACV 482

DB 423 YDEDEHYVVDYKELIKYKGSQVAPAELEEILLNKPICIRDVAVVGIPLDEAGELPSAFV 482

QY 483 VLESKTLTEKEVQDPIAAQVTPTKHLRGVVFVDSIPKPGTKGLIRKELREIFAQRAK 542

DB 483 VIQPGKEITAKEVYDIAERVSHTKYLRGVRFVDSIPRNVTKITRKELLKOLLEKSSK 542

RESULT 7

Protein-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) OR [validated] - lum
N:Alternate names: firefly-type luciferase
C:Species: Pyrophorus plagiophthalmus
C:Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
C:Accession: S29355

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence
A:Reference number: S29352; MUID:89242142

A:Accession: S29355

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 49.4%; Score 1388; DB 2; Length 543;

Best Local Similarity 48.7%; Pred. No. 5.1e-91;

Matches 263; Conservative 114; Mismatches 161; Indels 2; Gaps 1;

QY 3 EENIRHGERPRDIVHPGSAQOOLYQSLYKFASFPEALIDAHTNEVISYQAIFETSCRLAV 62

DB 5 ERNVIYGPPLPLEDKTAGEMLFRAIRKHSHPQAIQVDFGDESLSYKEFFETACLLAO 64

QY 63 STEQYGLNENNVVGCSENNINFPNVAALYLGIPVATSDMYTDGELTGHNLISKPTI 122

DB 65 SLHNCGYKNDVVSICAENKREFPIIAAWYIGMIVAPVNESYIPDELCKYMGISKPOI 124

QY 123 MFSSKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARYTDHTFDPLSTPKD 182

DB 125 VFCTKNILNKVLEVSQRTNFIKRIIILDTVENIHGCESLPNFISRYSDGNI--ANFKPLH 182

QY 183 FQPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPITVYGRTPVTPOTSILSLVPPFH 242

DB 183 YDPVEOVAALICSSGTGLPKGMVQTHQICVRLIHAIIDPEAGTQIIPGVTVLVVYPPFH 242

QY 243 AFGMFTTSLYFVVGKLVMLKFEAGFLTKTQNYKIPTIVVAPPVVMFLAKSPLVDQYD 302

DB 243 AFGESINLGYFVGLRVIMLRREFQEAFLKAIQDYEVRSIVNVVPAIILFLSKSPLVDKYD 302

QY 303 LSLTEVATGGAPLCKDVAEAVAKRLKLPGLIIOGYGLTETCCAVMTTPHNAVKTGSTRP 362

DB 303 LSSRELCCGAAPLAKAEVAEIAVKRLNLPICRGFGLTESTSANIHSLGDEFKSGSLGRV 362

QY 363 LPYIKAKVLNDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTIDKGMHLSGDIGY 422

DB 363 TPLMAAKIADRTGKALGPVQGEICIKGPMVSKGVVNNVATKEAIDDDGMHLSGDFGY 422

QY 423 YDEGNFFITVDRLKELIKYQVAPAELENLLQHPISADAGVTGVPDFGGQLPAACV 482

DB 423 YDEDEHYVVDYKELIKYKGSQVAPAELEEILLNKPICIRDVAVVGIPLDEAGELPSAFV 482

QY 483 VLESKTLTEKEVQDPIAAQVTPTKHLRGVVFVDSIPKPGTKGLIRKELREIFAQRAK 542

DB 483 VIQPGKEITAKEVYDIAERVSHTKYLRGVRFVDSIPRNVTKITRKELLKOLLEKSSK 542

RESULT 8

S29352

Protein-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) GR [validated] -

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000

C:Accession: S29352

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29352

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 49.1%; Score 1382; DB 2; Length 543;

[illegible][illegible]

RESULT 13
T07908
4-coumarate--CoA ligase (EC 6.2.1.12) 2 - western balsam poplar x cottonwood
C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
C:Accession: T07908
R:Allina, S.M.; Pri-Hadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.
Plant Physiol. 116, 743-754, 1998
A:Title: 4-Coumarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes, c
A:Reference number: Z16208; MUID:98150279
A:Accession: T07908
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-548 <ALL>
A:Cross-references: EMBL:AF008183; NID:g2911796; PIDN:AA039365.1; PID:g2911797
C:Genetics:
A:Gene: 4CL2
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:74-533/Domain: acetate--CoA ligase homology <ACL>

Query Match 30.1%; Score 847; DB 2; Length 548;
Best Local Similarity 36.8%; Pred. No. 1.8e-52;
Matches 196; Conservative 104; Mismatches 177; Indels 56; Gaps 15;

Qy 26 YQSLYKASFPPEALIDAHTNEVISAQIFETSCRLAVSIEQYGLNENNVGVCSENNINP 85
Db 33 FEKLSQKNDP-CLINGPTGDIYADVELTSRVSASGLYKLGQOGDVIILLQNSPEF 91
Qy 86 FNPVLAALYLGPVATSDMYTNDGELTGHNLISKPTIMFSSKKALPLTLRVOONLSFIKK 145
Db 92 VFALGASFGAISTANPFTTSRIKQATASAKLIITHAAVQQAQNDHV-K 150
Qy 146 VVVIDMYDINGVECVSTFVARYDHTDPD-LSPTPKDFDPLEKIALIMSSSGTGLPKG 204
Db 151 IMTIDSLTE---NCLHFSELTSSENEIPTVKIKPD-----IMALPYSSGTGLPKG 200
Qy 205 VVLSHRSLSLIR-----FVHSRDPYIGTRTPVQTSILSLVPPHAFGFTLSY 252
Db 201 VMLTHKGLVTSVAQVQDGENENLYFHERD-----VILCVPLFH---IYLSNV 246
Qy 253 FVVGK----VVMKKFEGALFKTIQNYKIPTIVVAPVPMVFLAKSPLDQYDLSLITE 308
Db 247 FLCGLRAGSAILVMOKFDVSLMDLVOKYVTIAPLVPPICIAKSPVVDYDLSLIRT 306
Qy 309 VATGAPLGKDVAAVAKRLKLP--IIQGYGLTET-----CCAVMITPHNAVKTGSTG 360
Db 307 VLSGAAPLGKEDTV--RAKLPNAKLGQGYGTEAGPVIAMCLAFAPKEPE-IKSGACG 363
Qy 361 RPLPYIKAKVLNDATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKGMHLHSGDI 420
Db 364 TVWRNAEMKIIVDPETGESOPRNTKTEICIRGQIMKGYLNDPEATERTIDKGMHLTGTDI 423
Qy 421 GYDDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTCVPDFGGQLPAA 480
Db 424 GYDDED-ELFIVDRKLKELIKYGVQVAPAELEAMLIAPHNIDSDAAVPMDEAGEVPA 482
Qy 481 CWVLESKLTLEKEVQDFIAAQVPTPKHLRGVGVFVDSIPKGTGKLRKELR 533
Db 483 FVVRNSGSKTTEDEIKQVSKQVIFYKRI-GRVFFTEAIPRPSGKILRKDLR 534

RESULT 14
B39827
4-coumarate--CoA ligase (EC 6.2.1.12) 2a - potato
N:Alternate names: 4-coumaroyl-CoA synthetase
C:Species: Solanum tuberosum (potato)
C:Date: 20-Mar-1992 #sequence_revision 20-Aug-1994 #text_change 05-May-2000
C:Accession: B39827
R:Becker-Andre, M.; Schulze-Lefert, P.; Hahlbrock, K.
J. Biol. Chem. 266, 8551-8559, 1991
A:Title: Structural comparison, modes of expression, and putative cis-acting elements of

A:Reference number: A39827; MUID:91217100
A:Accession: B39827
A:Molecule type: DNA
A:Residues: 1-545 <BDC>
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:76-535/Domain: acetate--CoA ligase homology <ACL>

Query Match 29.8%; Score 838; DB 1; Length 545;
Best Local Similarity 36.4%; Pred. No. 7.6e-52;
Matches 195; Conservative 109; Mismatches 171; Indels 60; Gaps 15;

Qy 26 YQSLYKASFPPEALIDAHTNEVISAQIFETSCRLAVSIEQYGLNENNVGVCSENNINP 85
Db 35 FENLSEFNSRP-CLIDGANDRIYTYAEVELTSRVAVGLNKLGIQOQDPTIMILLPNCPEF 93
Qy 86 FNPVLAALYLGPVATSDMYTNDGELTGHNLISKPTIMFSSKKALPLTLRVOONLS 140
Db 94 VFALGASFGAISTANPFTPAEVVQAKASSAKIVITOCAPAGVK---DYAIEDNL 150
Qy 141 SFIKKVVVIDMYDINGVECVS-TFVARYDHTDPDPLSFTPKDFDPLEKIALIMSSSGTT 199
Db 151 ----KVICVDSAP-----GCVFHSELIQSDHEIHPDVKIQPD-----VVALPYSSGTT 196
Qy 200 GLPGVVLHRSLSLIR-----FVHSRDPYIGTRTPVQTSILSLVPPHAFGME 247
Db 197 GLPGVVLHRSLSLIR-----FVHSRDPYIGTRTPVQTSILSLVPPHAFGME 245
Qy 248 TTLGVFV-VGLKVVMLKFEKALFKTIQNYKIPTIVVAPVPMVFLAKSPLDQYDLSL 306
Db 246 SVLLCALRVGAAILIMOKFDIAQFLKPKHVTIGTPVPIVLAIAKSPLVHNYDLSV 305
Qy 307 TEVATGGAPLGKDVAAVAKRLKLP--IIQGYGLTET-----CCAVMITPHNAVKTGS 358
Db 306 RTVMSGAAPLGKEDAV--RAKFPNAKLGQGYGTEAGPVLAMCLAFAPKEPD-IKSGA 362
Qy 359 TGRPLPYIKAKVLNDATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKGMHLHSG 418
Db 363 CGTVWRNAEMKIIVDPETGCSLPRNQPGEICIRGQIMKGYLNDPEATARTIEKGMHLTG 422
Qy 419 DIGYDDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTCVPDFGGQLP 478
Db 423 DIGFIDDDDELFIIVDRKLKELIKYGVQVAPAELEALLNHPDISDAAVPMDEAGEV 482
Qy 479 AACVLESKLTLEKEVQDFIAAQVPTPKHLRGVGVFVDSIPKGTGKLRKELR 533
Db 483 VAFVVRNSGSKTTEDEIKQVSKQVIFYKRIK-RVFFETVTPKSPSGKILRKDLR 536

RESULT 15
JU0311
4-coumarate--CoA ligase (EC 6.2.1.12) - rice
N:Alternate names: 4-coumaroyl-CoA synthetase
C:Species: Oryza sativa (rice)
C:Date: 30-Jun-1992 #sequence_revision 20-Aug-1994 #text_change 05-May-2000
C:Accession: JU0311
R:Zhao, Y.; Kung, S.D.; Dube, S.K.
Nucleic Acids Res. 18, 6144, 1990
A:Title: Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL.1.
A:Reference number: JU0311; MUID:91045096
A:Accession: JU0311
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-563 <ZHA>
A:Cross-references: EMBL:X52623; NID:g20160; PIDN:CAA36850.1; PID:g20161
A:Experimental source: strain Japonica
C:Genetics:
A:Gene: 4-CL.1
A:Introns: 351/3; 418/1; 489/2; 520/3
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:87-549/Domain: acetate--CoA ligase homology <ACL>

Query Match		29.6%	Score 833.5;	DB 1;	Length 563;
Best Local Similarity		39.3%	Pred. No. 1.7e-51;		
Matches 207;		Conservative 89;	Mismatches 192;	Indels 39;	Gaps 16;
Qy	30	YKFAFPPE-----AIIDAHTEVISAQIFETSCRLAVSIEQ--YGLNENNNVGVCSENN	82		
Db	42	YCFERLPEVAARPCLLIDGATGVLTADVDRLSRRLAALRAPLGRGGVVMSSLRNS	101		
Qy	83	INFFNPVLAALYLGIPVATSDMTDGLTGHNLNISKPTIMFSSKKALPLILRVQOQLSF	142		
Db	102	PEVLSFFAASRGAATVATPMSTPHIESQLAAAGATVVVITESMAAD---KLPSHSHG	158		
Qy	143	IKKVVVIDSMYDINGVECVSTFVARYTDHTDPLSFTPKD--FDPLEKIALIMSSSGTT	199		
Db	159	ALTVVVIDERD---GCLHFWDDLMSDEDEASPLAGDEDEKVFDPDDVVAL-PYSSGTT	213		
Qy	200	GLPKGVVLSHRSLT---IRFVHSRDPYIGTRTVPQTSILSLVPFHAFGMFTTISYFV-V	255		
Db	214	GLPKGVMLTHRSLSSTVSAQQVDGENPNIGLHA--GDVILCALPMFHIYSLNTIMCGLRV	271		
Qy	256	GLKVYMLKKFEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPLYDOYDLSSLTEVATGAP	315		
Db	272	GAAIVMRRFDLAAMMDLVERHRTIAPLPPIVVAVAKSEAAAARDLSSVRMVLGAAP	331		
Qy	316	LKQVAAEAVAKRLKLPGLI--QGYGLTET-----CCAVMITPHNAVKTGSTRPLPYIK	367		
Db	332	MKGIEDAF--MAKLPGAVLQOGYGMTAGPVLNCLAFAPKEPK-VKSGACGTVVRNAE	388		
Qy	368	AKVLDNATGKALGPERGEICFQSEMIMKGYNNNPEATIDTIDKDWLHSGDIGYDEDG	427		
Db	389	LKIIDPDGKSLGRNLRGEICIRGQQIMKGYLNNPEATKNTIDAEGLHTGDIGYVDDDD	448		
Qy	428	NFFIVDLRLKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFQGLPAACVWLESG	487		
Db	449	EIFIVDLRLKEIKYRGFOVAPAELEALLNTHPSTADAAVAVGL--KF-GEIPVAFVAKTEG	505		
Qy	488	KTLTEKEVODFIAAQVTPTKHLRGVWFVDSIPKGPCTGKLIRKELRE	534		
Db	506	SELSEDDVKQFVAKEVIYKKIR-EVFFVDKIPKAPSGKILRKELRK	551		

Search completed: September 4, 2002, 15:04:23
Job time: 142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:02:46 ; Search time 39.58 Seconds
(without alignments)
386.478 Million cell updates/sec

Title: US-09-993-874-2
Perfect score: 2812
Sequence: 1 MEEENIRHGERPRDIVHPSG.....KLIRKELREIFAORAPKSKL 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 763338 seqs, 245939087 residues

Total number of hits satisfying chosen parameters: 763338

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2812	100.0	545	US-09-791-537-100770 Sequence 100770,
2	2099.5	74.7	546	US-09-791-537-100774 Sequence 100774,
3	1584.5	56.3	550	US-09-791-537-123865 Sequence 123865,
4	1584.5	56.3	1172	US-10-122-706-4 Sequence 4, Appl1
5	1584	56.3	549	US-09-791-537-101683 Sequence 101683,
6	1583.5	56.3	550	US-09-791-537-8343 Sequence 8343, Ap
7	1583.5	56.3	552	US-10-111-723-2 Sequence 2, Appl1
8	1568.5	55.8	550	US-09-791-537-124509 Sequence 124509,
9	1568	55.8	551	US-09-791-537-124513 Sequence 124513,
10	1568	55.8	552	US-09-791-537-8770 Sequence 8770, Ap
11	1568	55.8	553	US-09-791-537-38560 Sequence 38560, A
12	1557.5	55.4	550	US-09-791-537-8344 Sequence 8344, Ap
13	1534	54.6	547	US-09-791-537-33330 Sequence 33330, A
14	1519	54.0	552	US-09-791-537-131812 Sequence 131812,
15	1516	53.9	666	US-09-581-894A-16 Sequence 16, Appl1
16	1516	53.9	715	US-09-581-894A-5 Sequence 5, Appl1
17	1515	53.9	552	US-09-791-537-131813 Sequence 131813,
18	1515	53.9	683	US-09-581-894A-14 Sequence 14, Appl1
19	1515	53.9	684	US-09-581-894A-18 Sequence 18, Appl1
20	1515	53.9	684	US-09-581-894A-20 Sequence 20, Appl1
21	1511.5	53.8	548	US-09-791-537-27875 Sequence 27875, A
22	1491.5	53.0	548	US-09-791-537-123864 Sequence 123864,
23	1485	52.8	548	US-09-791-537-45469 Sequence 45469, A
24	1483	52.7	548	US-09-791-537-53113 Sequence 53113, A
25	1458.5	51.9	548	US-09-791-537-73452 Sequence 73452, A
26	1457.5	51.8	548	US-09-791-537-17764 Sequence 17764, A

ALIGNMENTS

RESULT 1

US-09-791-537-100770
; Sequence 100770, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bloncomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 100770

; LENGTH: 545

; TYPE: PRT

; ORGANISM: Phrixothrix vivianii

US-09-791-537-100770

Query Match 100.0%; Score 2812; DB 5; Length 545;

Best Local Similarity 100.0%; Pred. No. 4.7e-235;

Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEEENIRHGERPRDIVHPSGAGQQLYQSLYKFPASPEAIIDAHTNEVISYAIQIFETSRL 60

Db 1 MEEENIRHGERPRDIVHPSGAGQQLYQSLYKFPASPEAIIDAHTNEVISYAIQIFETSRL 60

Qy 61 AVSISQYGLNENNVGVCSENNINFPVLAALYIGIPVATSDMYTDGELTGHNLISRP 120

Db 61 AVSISQYGLNENNVGVCSENNINFPVLAALYIGIPVATSDMYTDGELTGHNLISRP 120

Qy 121 TFMFSKKALPILRVOQNLSFIKKVWVDSMDYDINGVECVSTFVARYTDHFTDPLSFTP 180

Db 121 TFMFSKKALPILRVOQNLSFIKKVWVDSMDYDINGVECVSTFVARYTDHFTDPLSFTP 180

Qy 181 KDFDPLEKIALIMSSGTTGLPKGVVLSHRSITIRFVHSRDPYIGTRTPVQTSILSLVPF 240

Db 181 KDFDPLEKIALIMSSGTTGLPKGVVLSHRSITIRFVHSRDPYIGTRTPVQTSILSLVPF 240

Qy 241 HFAFCMFTLSFVVGKLVNKKFEGALFKTIQNKIPTIIVAPPVWVFLAKSLPVDQ 300

Db 241 HFAFCMFTLSFVVGKLVNKKFEGALFKTIQNKIPTIIVAPPVWVFLAKSLPVDQ 300

Qy 301 YDLSSITVEATGAPLGKDVAAEAVAKRLKLPGLIQYGLTETCCAVMITPHNAVKTGSTG 360

Db 301 YDLSSITVEATGAPLGKDVAAEAVAKRLKLPGLIQYGLTETCCAVMITPHNAVKTGSTG 360

27 1454 51.7 545 5 US-09-791-537-61684 Sequence 61684, A
28 1453.5 51.7 548 5 US-09-791-537-123832 Sequence 123832, A
29 1393 49.5 543 5 US-09-791-537-60969 Sequence 60969, A
30 1388 49.4 543 5 US-09-791-537-118498 Sequence 118498, A
31 1382 49.1 543 5 US-09-791-537-60967 Sequence 60967, A
32 1377 49.0 543 5 US-09-791-537-56394 Sequence 56394, A
33 1377 49.0 544 5 US-09-791-537-8822 Sequence 8822, Ap
34 864.5 30.7 542 5 US-09-791-537-49392 Sequence 49392, A
35 862 30.7 547 5 US-09-791-537-19666 Sequence 19666, A
36 855 30.4 542 5 US-09-791-537-120527 Sequence 120527, A
37 847.5 30.1 636 5 US-09-791-537-84582 Sequence 84582, A
38 847 30.1 545 5 US-09-791-537-107375 Sequence 107375, A
39 847 30.1 548 5 US-09-791-537-107378 Sequence 107378, A
40 838 29.8 545 5 US-09-791-537-69014 Sequence 69014, A
41 836.5 29.7 553 5 US-09-791-537-56966 Sequence 56966, A
42 834 29.7 545 5 US-09-791-537-96123 Sequence 96123, A
43 833.5 29.6 563 5 US-09-791-537-96123 Sequence 96123, A
44 829.5 29.5 544 6 US-10-202-307-349 Sequence 349, App
45 823.5 29.3 557 5 US-09-791-537-120554 Sequence 120554, A

Qy 361 RPLPYIKARVLDNATGKALGPCERGEICFQSEMIMKGYNNPEATIDTDKDGWLHSGDI 420
Db 361 RPLPYIKAKVLDNATGKALGPCERGEICFQSEMIMKGYNNPEATIDTDKDGWLHSGDI 420
Qy 421 GYYDEGNGFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFQGLPAA 480
Db 421 GYYDEGNGFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFQGLPAA 480
Qy 481 CWVLESGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIFAORA 540
Db 481 CWVLESGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIFAORA 540
Qy 541 PKSKL 545
Db 541 PKSKL 545

RESULT 2
US-09-791-537-100774
; Sequence 100774, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100774
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Phrixothrix hirtus
US-09-791-537-100774

Query Match 74.7%; Score 2099.5; DB 5; Length 546;
Best Local Similarity 71.1%; Pred. No. 2.8e-173; Mismatches 66; Indels 9; Gaps 3;
Matches 391; Conservative 84;

Qy 1 MEEENIRHGERPRDIVHPGSGAQOOLYQSLYKFPASPEAIIIDAHTNEVISAQIFETSCRL 60
Db 1 MEEENVNGDRPLVFPGTAGLQLYSLYKYSYITDGIIDAHTNEVISAQIFETSCRL 60
Qy 61 AVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSNMYTDCBELTGHNLNISK 120
Db 61 AVSLEKYGLDHNNVVAICSENNIHFFGPLIAALYQGI PMATSNMYTEREMIGHNLNISK 120
Qy 121 TMFSSKKALPLILRVQONLSFIKVVVIDSMYDINGVECVSTFVARVTDHTFDPLSPT 180
Db 121 CLMFCSKSLPFLKVKQHLDFLRVVIDSMYDINGVECVSTFSDRNTDIAFDVKNP 180
Qy 181 KDFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSILSLV 240
Db 181 KEFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSILAIAP 240
Qy 241 HHAGCMFTTLYFVVVGLKVMVKKFEGALFLKTONYKIPTIVVAPPVVMVFLAKSPVDQ 300
Db 241 HHAGLFTALAYFPVGLKIVMVKFEGEFLKTONYKIASIVPPPTMVYLAKEPLVDE 300
Qy 301 YDLSLSTEVATGGAPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMTTPHN-AVKTGST 359
Db 301 YNCSSLTEIASGSPGLGRDIADKVAKRLKVGILQGYGLTETCSALLISPNDRLEKKA 360
Qy 360 GRPLPYIKAKVLDNATGKALGPCERGEICFQSEMIMKGYNNPEATIDTDKDGWLHSGD 419
Db 361 GTPMPYQVKVIDINTGKALGPKEKGEICFQSEMIMKGYNNPEATIDTDKDGWLHSGD 420
Qy 420 IGYDDEGNGFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFQGL-- 477
Db 421 LCYDDEGNGFFIVDRUKELIKYGVQVAPAELENLLQHPNIDAGVTGVPDEFQGLV 476

Qy 478 --PAACVVLESKGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREI 535
Db 477 NYLSACVVLEPGKMTTEKEVODYIAELVTTTKHLRGVVFIDSIPKGTGKLMRNLRAI 536
Qy 536 FAQAPAKSKL 545
Db 537 FAREQAKSKL 546

RESULT 3
US-09-791-537-123865
; Sequence 123865, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 123865
; LENGTH: 550
; TYPE: PRT
; ORGANISM: pdb 1LCI
US-09-791-537-123865

Query Match 56.3%; Score 1584.5; DB 5; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.3e-128; Mismatches 137; Indels 5; Gaps 3;
Matches 305; Conservative 103;

Qy 1 MEE-ENIRHGERPRDIVHPGSGAQOOLYQSLYKFPASPE--AIIDAHTNEVISAQIFETS 57
Db 1 MEDAKNIKKGPAFPYPLEDGTAGEQLHKAKRYALVPGTIAFTDAHIEVNITYAEVFE 60
Qy 58 CRLAVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSNMYTDCBELTGHNL 117
Db 61 VRLAEMKRYGLNTHIRVWCSENSLQFMPVLGALFVGAVAPANDIYNRELLNSMNI 120
Qy 118 SKPTIMSSKKALPLILRVQONLSFIKVVVIDSMYDINGVECVSTFVARVTDHTFDPL 177
Db 121 SQTPTVFWYSKGLQKLVNQKLPKLPKLPKLPKLPKLPKLPKLPKLPKLPKLPKLP 180
Qy 178 FTKPDFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSIL 237
Db 181 FVPSFDRDKTIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSIL 240
Qy 238 VPFHAGCMFTTLYFVVVGLKVMVKKFEGALFLKTONYKIPTIVVAPPVVMVFLAKSP 297
Db 241 VPFHAGCMFTTLYFVVVGLKVMVKKFEGALFLKTONYKIPTIVVAPPVVMVFLAKSP 300
Qy 298 VDQYDLSLSTEVATGGAPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMTTPHNNAV 357
Db 301 IDKVDLSNLHLSAGCAPLSKEGVAEAVAKRFLPGIRGYGLTETTSAILITPEGDDRP 360
Qy 358 STGRPLPYIKAKVLDNATGKALGPCERGEICFQSEMIMKGYNNPEATIDTDKDGWLH 417
Db 361 AVGVVVPFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGVYNNPEATNALIDKDWLH 420
Qy 418 GDYGYDEGNGFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFQGL 477
Db 421 GDIAWDEDEHFFIVDRUKSLIKYGVQVAPAELESILLQHPNIPFDAGVAGLDDAGEL 480
Qy 478 PAACVVLESKGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIF- 536
Db 481 PAACVVLEHGTMTKEIYDVVASQVTTAKKLRGVVFVDEVPKGLTGKLDARKIREIL 540
Qy 537 -AQRAPAKSKL 545
Db 537 -AQRAPAKSKL 545

Db 541 KAKGGKSKL 550

RESULT 4

```

US-10-122-706-4
; Sequence 4, Application US/10122706
; GENERAL INFORMATION:
; APPLICANT: Srinivasan, Malthreyan
; APPLICANT: Refilser, Michael
; TITLE OF INVENTION: Sulfurylase-Luciferases
; FILE REFERENCE: 21465-504
; CURRENT APPLICATION NUMBER: US/10122,706
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/335,949
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-122-706-4

```

Query Match 56.3%; Score 1584.5; DB 6; Length 1172;

Best Local Similarity	55.5%	Pred. NO. 3.9e-128;
Matches 305; Conservative	103; Mismatches	137; Indels 5; Gaps 3;

Matches	305;	Conservative	103;	Mismatches	137;	Indels	5;	Gaps
Qy	1	MEZ-ENTRHGERPRD	VHPGSAGOOYLSYKFPASPE--AII	DAHTNEVSYAQIFETS	57			
Db	109	MEDAKNIKKGPAPFY	PLEDDTAGEOHLKAMKRYALVPGTIAFT	DAHIEVNITYAEVFEWS	168			
Qy	58	CRLAVSIEOYGLNENN	VVGVCSENNINFFNPVLAALYLGIPVATS	NDMYTDGELTGHLLNI	117			
Db	169	VRUAEAMKRYGLNTN	HKRVCSSENSQFMPVLGALFIGVAVAPANDI	YNERELLSMNI	228			
Qy	118	SKPTIFSSKKLPLIL	RVOONLSFKVVV	IDSMYDINGVECVSTFYARYTDHTFDPLS	177			
Db	229	SQPTVVFVSKKGLOK	ILNVOKLPILOKII	INDSKTDYOGFQSMYTFVTSHPLPGFNEVD	288			
Qy	178	FTPKDEDPLEKIAL	IMSSGTTGLPGVVLVSHRSLTIRFVHSRDP	YIGTRTVQPTSILSL	237			
Db	289	FVPESEDRDKTIAL	IMSSGTTGLPGKVALPHRTACVRFSHARD	PLFGNQLIIPDTAILSV	348			
Qy	238	VPRHHAFGMETTLSY	FVVGVLKVVLKFFEGALFLKTIQNKIPTI	VVAVPVMVFLAKSL	297			
Db	349	VPRHHGFGMETTIGY	LTCGPRVVLMYRFEELFLRSLQDYKIQSALL	VPTLFSFFAKSTL	408			
Qy	298	VOQYDLSLLEFVATG	GAPLCKDVAEAVAKRLKLP	IGIIOGYGLTETCCAYMITPHNAVKTG	357			
Db	409	IDKYDLSNLHIEIASG	NPLSKVEYGEAVAKRFLPGIRQGYGLTETTS	SALITPEGGDKPG	468			
Qy	358	STGRPLPYIKAKVLD	NATGKALGPGERGEICFOSEMIMKGYVNN	PEATIDTIDKDGWLHS	417			
Db	469	AVGVKVPVFFEAKE	VVDLDGTGTLGVNORGELCVRGPIMSGYVNN	PEATNALIDKDGWLHS	528			
Qy	418	GDIGYDEDGNFFIV	DRKELIKYKGVQVAPALENLLLOHFSIADAG	VTGVPDDEGGOL	477			
Db	529	GDITAWDEDHFFIV	DRKLSLIKYKGVQVAPALESILLOHPNIFDAG	VAGLUPDDDAGEL	588			
Qy	478	PAACVWLESCKTLE	KREVQDFIAAQVTPTKHLRGVVVFVDS	TPKPGPTGKILKELREIF-	536			
Db	589	PAAVVWLEHCKTTE	KREIVDYVASQVTTAKLRGGVVVFVDEVP	KGLTGKLDARKIREILL	648			
Qy	537	-AORAPKSKL	545					
Db	649	KAKKGKSKL	658					

RESULT

US-09-791-537-101683
; Sequence 101683, Application US/09791537
; GENERAL INFORMATION:

```

; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101683
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Reporter vector p2luc
US-09-791-537-101683

```

Query Match

Best Local Similarity	55.3%	Pred. NO. 1.4e-128;	
Matches 303:	Conservative	103:	Mismatches 138;
			Indels 4;
			Gaps 2;

Qy	2	EENIRIGERPRD	IVHPGSGAGQOYLQSLYKPA	SPPE--AIIDAHTNEVISAQIFET	SCR	59
Db	3					60
Db	2	DAKNIKKG	PAPFPLEDGTAGEQLHAKMKRYAV	PGPTIAFTDAHIEVNTIYA	FENSVR	61
Qy	60	LAYSIVG	YGLNENNVGVCSENINFPNPVLAALY	LGIPVATSDMKYTDGELTGH	LNISK	119
Db	62	LAEAMKRYG	LNTNHRIVVCSSENSLOFFMPVLGAL	FIGVAVAPANDIYNRELLNSMN	ISQ	121
Qy	120	PTIMSSFK	GLPLTLRVOQNLSEFIKKVWVSDMSY	DINGVECVSTFVARYDTH	TFDPLSFT	179
Db	122	PTVVVFSK	KGLOKLTINVOKKLP	PIIOKIIINDSKTDYOGFQSMYTFVT	SHLP	181
Qy	180	PKDPD	PLEKIALIMSSGGTGLPKGVVLSHRS	LUTIRPVHSRDPYIGRTRVPQTS	IIUSLVP	239
Db	182	PESFD	RKTTIALIMNSSGTGLPKGVALPHUR	TACVRESHARDPTFGNQLIPD	TAILSVWP	241
Qy	240	FHHAFG	FTTILSYFVWGLKVVMKKKPGCAL	FLKTIQNYKIPITVAVPMVAVFLAKS	PLVD	299
Db	242	FHHGFG	FTTILSYLIGCFRVLVWYRFEELFL	RSLODYKIQSALLVLTFLFSF	AKSTLID	301
Qy	300	QYDLS	STEATGAPJGUKDVAEAVAKRLKPG	ITIOQYGLITETCCAVMITPHNAVKT	GTST	359
Db	302	KYDLS	NLHETASGAPLSKEVGEAVARFHP	LGIRGOYGLTETTSALITPEGDDK	PGAV	361
Qy	360	GRPLPY	TAKVLDONATKALGPBERGETCFOSE	IMKGYNNPEATIDTIDKQGLW	HSGD	419
Db	362	GKVVP	PEAKVWDLDTCKTLGVNQRGELCVR	PGMISGYNNPEATNALIDKQGLW	HSGD	421
Qy	420	IGYDE	DCNFFIVDRKLKELTKYGYQVAPAE	LENLLLOHPSIADAGVTVGPDEFGG	OLPA	479
Db	422	IAYWDE	DEHFFIVDRKLSLTKYGYQVAPAE	LESILLOHFNIFDAGVAGLPDDD	DAGELPA	481
Qy	480	ACVVL	ESGKTLTEKEVODFTAAQVPTPKHL	RGGVVFVDSIPKGTPIRKLEIRE	IE--A	533
Db	482	AVVLE	HGRKWTKEIIVDYVASQVTTAKKL	RGGVVFVDEVPKGLTGLDK	ARKIREILKA	541
Qy	538	QRAPK	SKL	545		
Db	542	KKGK	SKL	549		

RESULT 6

US-09-791-537-8343
; Sequence 8343. Application US/09791537

? GENERAL INFORMATION: 057/09791537
 ? APPLICANT: Blomox, Inc.
 ? APPLICANT: Debe, Derek
 ? APPLICANT: Danzer, Joseph
 ? TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURE
 ? TITLE OF INVENTION: METHODS OF USE THEREOF
 ? FILE REFERENCE: 261/210
 ? CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8343
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Photinus pyralis
US-09-791-537-8343

Query Match 56.3%; Score 1583.5; DB 5; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.6e-128;
Matches 305; Conservative 102; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGRPRDIVHPSAGQOLYOSLYKFAFPF--AIIAHTNEVTSYAQIFETS 57
Db 1 MEDAKNIKKGPAPYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMS 60
Qy 58 CRLAVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNOMYTDELTGHLNI 117
Db 61 VRLAEAMKRYGLNTHRIWVCSNSLOQFMPVLGALFIGVAVAPANDIYNRELLNSNI 120
Qy 118 SKPTIMSSKKALPLILRVQONLSFIKKVVVDSMYDINGVECVSTFVARYTDHTFDPLS 177
Db 121 SQTVPVFSKKGKLOKILNVQKKLPYIIQKIIIMDSKTDYQGFQSMYTFVTSHPGCFNEYD 180
Qy 178 FTKDFDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVPOTSILSL 237
Db 181 FVPESFDRDKTIALIMSSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240
Qy 238 VPFHAFGCMFTLSYFVVGVLKVMKKEGALFLKTIQNYKIPTIVVAPPVYVFLAKSPL 297
Db 241 VPFHFGCMFTLGLYICGFRVLMYRFEELFLRSLODYKIQSALLVPTLFSFFAKSTL 300
Qy 298 VDQYDLSLSEVATGGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRPHLPGRQGYGLTETTSAILITPKGDKPG 360
Qy 358 STGRPLPIYAKVLDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTDKDGWLHS 417
Db 361 GVGKVPVFEAKVVDLTGKTLGVNQRGELCVRGPMIMSGVYNNPEATNALIDKDGWLHS 420
Qy 418 GDIGYDEGDNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTCVDFEGQOL 477
Db 421 GDIAWYDEDEHFFIVDRUKLSIKYGVQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480
Qy 478 PAACVWLESGLTTEKEVQDFIAAQVPTKHLRGVGVFVDSIPKGPCKLIRKELREIF- 536
Db 481 PAAVVLHGKMTKEIYDVIVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIRELI 540
Qy 537 -AORAPSKL 545
Db 541 KAKKGKSKL 550

RESULT 7

US-10-111-723-2
; Sequence 2, Application US/10111723
; GENERAL INFORMATION:
; APPLICANT: The Secretary of State for Defence
; APPLICANT: White, Peter J
; APPLICANT: Willey, Tara L
; APPLICANT: Price, Rachel L
; APPLICANT: Murphy, Melanie J
; APPLICANT: Squirrell, David
; TITLE OF INVENTION: Novel Enzyme
; FILE REFERENCE: DERA/IPD/PI247/MOD
; CURRENT APPLICATION NUMBER: US/10/111.723
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: GB 9925161.3
; PRIOR FILING DATE: 1999-10-26
; PRIOR APPLICATION NUMBER: GB 0016744.5
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Encoded amino acid
; OTHER INFORMATION: sequence of the cDNA sequence of synthetic
; OTHER INFORMATION: luciferase gene
US-10-111-723-2

Query Match 56.3%; Score 1583.5; DB 6; Length 552;
Best Local Similarity 55.5%; Pred. No. 1.6e-128;
Matches 305; Conservative 102; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGRPRDIVHPSAGQOLYOSLYKFAFPF--AIIAHTNEVTSYAQIFETS 57
Db 3 MEDAKNIKKGPAPYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMS 62
Qy 58 CRLAVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNOMYTDELTGHLNI 117
Db 63 VRLAEAMKRYGLNTHRIWVCSNSLOQFMPVLGALFIGVAVAPANDIYNRELLNSNI 122
Qy 118 SKPTIMSSKKALPLILRVQONLSFIKKVVVDSMYDINGVECVSTFVARYTDHTFDPLS 177
Db 123 SQTVPVFSKKGKLOKILNVQKKLPYIIQKIIIMDSKTDYQGFQSMYTFVTSHPGCFNEYD 182
Qy 178 FTKDFDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVPOTSILSL 237
Db 183 FVPESFDRDKTIALIMSSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 242
Qy 238 VPFHAFGCMFTLSYFVVGVLKVMKKEGALFLKTIQNYKIPTIVVAPPVYVFLAKSPL 297
Db 243 VPFHFGCMFTLGLYICGFRVLMYRFEELFLRSLODYKIQSALLVPTLFSFFAKSTL 302
Qy 298 VDQYDLSLSEVATGGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 303 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRPHLPGRQGYGLTETTSAILITPKGDKPG 362
Qy 358 STGRPLPIYAKVLDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTDKDGWLHS 417
Db 363 AVGVVPEFEAKVVDLTGKTLGVNQRGELCVRGPMIMSGVYNNPEATNALIDKDGWLHS 422
Qy 418 GDIGYDEGDNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTCVDFEGQOL 477
Db 423 GDIAWYDEDEHFFIVDRUKLSIKYGVQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 482
Qy 478 PAACVWLESGLTTEKEVQDFIAAQVPTKHLRGVGVFVDSIPKGPCKLIRKELREIF- 536
Db 483 PAAVVLHGKMTKEIYDVIVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIRELI 542
Qy 537 -AORAPSKL 545
Db 543 KAKKGKSKL 552

RESULT 8

US-09-791-537-124509
; Sequence 124509, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 124509

Db 241 VPFHGFCHFTTLGVLICGRVVLHYRFEELFLRLSLQDYKIQSALLVPTLESFAKSTL 300
QY 298 VDOYDLSSLETEVATGGAPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGAPLSKEVGEAVAKRFLPLGIRQYGLTETTSAILITPEGDDKPG 360
QY 358 STGRPLPYTKAKVLONATKALGPGERGICFQSEMIMKGYNNPEATIDTIDKDGWLHS 417
Db 361 AVGVVPPFEAKVVDLDTGTLGNQRGELCVGRGPMIMSGYNDPEATNALIDKDGWLHS 420
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPISIADAGVTGVPDEFGGOL 477
Db 421 GDIATWDEHEFFIVDRKLSLKIYGCQVAPAELESILLOHPNIFDAGVAGLPGDAGEL 480
QY 478 PAACVWLESGTKLTKKEVODFAAQTPTKHLRGVGVFVDSIPKPGTGLKIRKELREIF- 536
Db 481 PAAVVVLEHGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREIL 540
QY 537 -AORAPKSKL 545
Db 541 KAKKGKSKL 550

RESULT 13
US-09-791-537-33330
; Sequence 33330, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33330
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Lampyris noctiluca
US-09-791-537-33330

Query Match 54.6%; Score 1534; DB 5; Length 547;
Best Local Similarity 53.5%; Pred. No. 3.1e-124;
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;

QY 1 MEE-ENIRGERPRDIVHPGSGAQOYLQSLYKFAFPE--AIIDAHTNEVSYAQIFETS 57
Db 1 MEDAKNIMHGPAFFYPLEDGTAGEQLHAKKRYAOPGTIAFTDAHAENVITYSEYFEMA 60
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSNDMVTDGELTCHLNI 117
Db 61 CRLAETNKRYGLGLQHIIAVCSENSLOFPMPVCGALFIGVGVASTNDIYNRELNSLSI 120
QY 118 SKPTIMFSSKALPLILRVOQNLSPFKKVVVIDSMDYINGVCEVSTFVARYTDHTFDPLS 177
Db 121 SQTPIVSCSKRALQKILGVOKKPLPIQKIVILDSREDYMGKQSMYSFIESHLPAAGNEYD 180
QY 178 FYPKDFDPLEKTALINSSSGTGLPGVVLVSHRSLTIRVHSDRDIYGTTRTPQTSILS 237
Db 181 YTPDSFRETATATALINSSSGTGLPGVVLVSHRSLTIRVHSDRDIYGTTRTPQTSILS 240
QY 238 VPFHAFGFTTSLYFVVGKLVYMLKKFEGALFKTIQYKTIPTIVAPPVMMFLAKSP 297
Db 241 IFPHFGFCHFTTLGVLICGRVVLHYRFEELFLRLSLQDYKIQSALLVPTLESFAKSTL 300
QY 298 VDOYDLSSLETEVATGGAPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 357
Db 301 VDKYDLSNLHEIASGAPLSKEVGEAVAKRFLPLGIRQYGLTETTSAILITPEGDDKPG 360
QY 358 STGRPLPYTKAKVLONATKALGPGERGICFQSEMIMKGYNNPEATIDTIDKDGWLHS 417

Db 361 ACQKVVPPFSAKIVDLDTGTLGNQRGELCVGRGPMIMKGYNNPEATNALIDKDGWLHS 420
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPISIADAGVTGVPDEFGGOL 477
Db 421 GDIATWDEHEFFIVDRKLSLKIYGCQVAPAELESILLOHPNIFDAGVAGLPGDAGEL 480
QY 478 PAACVWLESGTKLTKKEVODFAAQTPTKHLRGVGVFVDSIPKPGTGLKIRKELREIFA 537
Db 481 PAAVVVLEHGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREIL- 539
QY 538 QORAPKSKL 545
Db 540 MMGKSKL 547

RESULT 14
US-09-791-537-131812
; Sequence 131812, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 131812
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Photuris pennsylvanica
US-09-791-537-131812

Query Match 54.0%; Score 1519; DB 5; Length 552;
Best Local Similarity 52.8%; Pred. No. 6.2e-123;
Matches 287; Conservative 102; Mismatches 151; Indels 4; Gaps 3;

QY 3 EENIRGERPRDIVHPGSGAQOYLQSLYKFAFPE--AIIDAHTNEVSYAQIFETS 60
Db 4 EENILGPPPPYPLEEGTAGELHRAITRYAAVPGTLAYTDVHTLEVITYKEFLDVTCL 63
QY 61 AVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSNDMVTDGELTCHLISK 120
Db 64 AEMKNYGLGLQHTISVCSENCVQFPMPVCAALYIGVATAPINDIYNRELNSLSISQ 123
QY 121 TMFSSKALPLILRVOQNLSPFKKVVVIDSMDYINGVCEVSTFVARYTDHTFDPLSFT 180
Db 124 TVVFTSRNSLQKILGVQSLPVKIKIIMLDTKDYLGYSQSMOSFMKEHPANFNVSAP 183
QY 181 KDFDPLEKTALINSSSGTGLPGVVLVSHRSLTIRVHSDRDIYGTTRTPQTSILS 240
Db 184 LSFD-LDRVACIMNSSSGTGLPGVVLVSHRSLTIRVHSDRDIYGTTRTPQTSILS 242
QY 241 HIAFGFTTSLYFVVGKLVYMLKKFEGALFKTIQYKTIPTIVAPPVMMFLAKSP 300
Db 243 HIAFGFTTSLYFVVGKLVYMLKKFEGALFKTIQYKTIPTIVAPPVMMFLAKSP 302
QY 301 YDLSSLETEVATGGAPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 360
Db 303 YDLSSLETEVATGGAPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 362
QY 361 RPLPYTKAKVLONATKALGPGERGICFQSEMIMKGYNNPEATIDTIDKDGWLHS 420
Db 363 KYVPFSLKVLDTLNTGKLGPNRGEICFTGPMIMKGYNNPEATREIDEDEGHIHSGDI 422
QY 421 GYDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPISIADAGVTGVPDEFGGOLPAA 480
Db 423 GYDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPISIADAGVTGVPDEFGGOLPAA 482

Qy 481 CVMLESGKLTKEVQDFTAAQVPTKHLRGVVFVDSIPKPTGKLIIRKELREIFAORA 540
Db 483 VVVLREGSKSITEKEIQDYVAGOVTSKKLRGGVEFVKPVGFTGKIDTRKIKEILI-KA 541
Qy 541 PKSK 544
Db 542 QKGK 545
RESULT 15
US-09-581-894A-16
; Sequence 16, Application US/09581894A
; GENERAL INFORMATION:
; APPLICANT: Fong, Tung Ming
; APPLICANT: van der Ploeg, Leonardus H. T.
; APPLICANT: Tota, Michael R.
; TITLE OF INVENTION: C-TERMINAL REGION OF AGOUTI-RELATED
; TITLE OF INVENTION: TRANSCRIPT (ART) PROTEIN
; FILE REFERENCE: 20146P
; CURRENT APPLICATION NUMBER: US/09/581.894A
; CURRENT FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: PCT/US98/26457
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/069,747
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Human
US-09-581-894A-16

Query Match 53.9%; Score 1516; DB 5; Length 666;
Best Local Similarity 51.4%; Pred. No. 1.5e-122;
Matches 286; Conservative 105; Mismatches 151; Indels 14; Gaps 3;
Qy 3 EENIRHGERPRDIVHPGAGQOLYQVLYKFAFPE--AIIDAHTNEVISYAOIFETSCRL 60
Db 95 ENNIIIGPPPPYPLEGTAGEQLHRAISRYAAVPGTLAYTDVHTTELEVITYREFLDVTCRL 154
Qy 61 AVSIEOYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSNMDYTGELTGHNLNISKP 120
Db 155 AEAMKNYGLQRTISVCSENCVPQFMPICAAALYGVVATPTNDIYNRELYNSLSISQP 214
Qy 121 TIFSSKKALPLILRVQONLSFIKKVWIDSNVDINGVECYSTFFVARYTDHTFDPLSFTP 180
Db 215 TVVFTSRNSLQKILGVQSRPLPIIKIILDGKKDYLGYOSMOSFMKEHVPANFNVSAPKP 274
Qy 181 KDFDPLEKIALIMSSGGTGLPKGVVLSHRSLSITIRFVHSRDPITYGTRTPQTSILSLVPF 240
Db 275 LSFD-LDRVACIMNSSGSTGLPKGVPISHRNTIYFESHCRDPVFCNQIIPDTTILCAVPF 333
Qy 241 HFAFGFTTSLFVYVGLKVMMLKKEGALFLKTIQNYKIPTIVVAPPVWFLAKSPLVDQ 300
Db 334 HFAFGFTNLGLYLGCGFHVLMYRNEHLFLQTLQDYKOSALLVPTVLAFIAKKNPLVDK 393
Qy 301 YDLSLSTEATCGAPLGRDVAEAVAKRLKPGIIQGYGLTETCCAVMITPHNAVKTGSTG 360
Db 394 YDLSNLHEIASGGAPLSKEISEIAAKRFLPGIRQGYGLTETTCATVITAEGEFKLGAVG 453
Qy 361 RPLPYIKARVLNATGKALGPERGEICFQSEMIKMGYNNPEATIDTIDKDWLHSGDI 420
Db 454 KVPFYSULKVLDLNTGKLGPNERGEICFKGPMIMKGYINNPEATRELIDEGWIHSGDI 513
Qy 421 GYDDEGNFFIVDRKELIKYGYQVPAELENLLQHPSIADAGVTGVPDEFGQLPAA 480
Db 514 GYFDEGHHYIVDRKLSLIKGYQVPAELENLLQHPFIEDAGVAGVDEVDGLPCA 573
Qy 481 CVMLESGKLTKEVQDFTAAQVPTKHLRGVVFVDSIPKPTGKLIIRKELREIFA--- 536
Db 574 VVVLREGSKSITEKEIQDYVAGOVTSKKLRGGVEFVKPVGFTGKIDTRKIKEILIRAQ 633

Qy 537 -----AORAPKSKL 545
Db 634 KGKSKAKLGGPEQKL 649

Search completed: September 4, 2002, 15:05:10
Job time: 144 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:00:26 ; Search time 34.33 Seconds
(without alignments)
1763.333 Million cell updates/sec

Title: US-09-993-874-2

Perfect score: 2812

Sequence: 1 MEEENTRHGERPRDIHVPGS.....KLIRKELREIFAQAPKSKL 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1588.5	56.5	550	16	AA866627 Firefly luciferase
2	1585.5	56.4	550	17	AA898515 P. pyralis variant
3	1584.5	56.3	550	15	AA850011 Firefly luciferase
4	1584.5	56.3	550	20	AA808523 P. pyralis pTetLuc
5	1584.5	56.3	550	22	AA831172 Amino acid sequenc
6	1583.5	56.3	550	16	AA83006 Firefly luciferase
7	1583.5	56.3	550	17	AA898518 P. pyralis variant
8	1583.5	56.3	552	22	AA835498 Synthetic lucifera
9	1581.5	56.2	815	13	AA828127 Lux::npt-II fusion
10	1577.5	56.1	743	15	AA854867 Hepatitis C Virus
11	1577.5	56.1	743	16	AA872801 Vaccinia virus vec

12	1576.5	56.1	550	20	AAW92747 P. pyralis luc pro
13	1568.5	55.8	550	18	AAW32936 Modified firefly l
14	1568.5	55.8	550	22	AA835055 Plant signal trans
15	1568	55.8	551	18	AAW32937 Modified firefly l
16	1524.5	54.2	552	20	AAW99364 Chimeric firefly l
17	1521.5	54.1	552	20	AAW99365 Chimeric firefly l
18	1516	53.9	666	20	AAW21871 Amino acid sequenc
19	1516	53.9	715	20	AAW21860 Amino acid sequenc
20	1515	53.9	552	16	AAW67007 Firefly luciferase
21	1515	53.9	683	20	AAW21869 Amino acid sequenc
22	1515	53.9	684	20	AAW21873 Amino acid sequenc
23	1515	53.9	684	20	AAW21875 Amino acid sequenc
24	1493.5	53.1	548	18	AAW27378 L. cruciata lucife
25	1492.5	53.1	548	18	AAW27376 L. cruciata lucife
26	1492.5	53.1	552	20	AAW99366 Chimeric firefly l
27	1491.5	53.0	548	10	AAW94367 Recombinant lucife
28	1491.5	53.0	548	11	AAW95788 Luciferase. Synth
29	1491.5	53.0	548	22	AAW51161 Modified Luciola c
30	1490.5	53.0	548	18	AAW27375 L. cruciata lucife
31	1488.5	52.9	548	18	AAW27377 L. cruciata lucife
32	1487.5	52.9	548	14	AAW30802 Thermostable lucif
33	1487.5	52.9	548	14	AAW33709 Luciferase mutant
34	1486.5	52.9	548	12	AAW14122 Luciferase mutant
35	1485.5	52.8	548	12	AAW14120 Luciferase mutant
36	1485.5	52.8	548	12	AAW14124 Luciferase mutant
37	1485	52.8	548	16	AAW84178 Luciferase of Hota
38	1484.5	52.8	548	12	AAW14119 Luciferase. Lucio
39	1483.5	52.8	548	12	AAW14123 Luciferase mutant
40	1483.5	52.8	548	12	AAW14125 Luciferase mutant
41	1482	52.7	544	22	AAU01229 Beetle luciferase
42	1481.5	52.7	548	12	AAW14121 Luciferase mutant
43	1477	52.5	544	22	AAU01232 Beetle luciferase
44	1475	52.5	544	22	AAU01228 Beetle luciferase
45	1475	52.5	546	20	AAW06859 Mutant luciferase

ALIGNMENTS

RESULT 1

AA866627

ID AA866627 standard; Protein; 550 AA.

AC AA866627;

DT 12-FEB-1996 (first entry)

DE Firefly luciferase (A215L,E354X) engineered mutant.

KW Luciferase; thermostability; enzyme stabilization; firefly;

KW Photinus pyralis; luc gene; site-directed mutagenesis; label;

KW ATP analysis.

OS Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 354

FT /note= "amino acid at position 354 is any amino acid other than Glu, pref. also not Gly, pro or Asp, and is pref. Trp, Val, Leu, Ile, Asn or esp. Arg or Lys, their analogues or modified forms"

XX WO9525798-A1.

XX 28-SEP-1995.

PF 22-MAR-1995; 95WO-GB00629.

XX 20-JAN-1995; 95GB-0001170.

PR 23-MAR-1994; 94GB-0005750.

XX (MNA) UK SEC FOR DEFENCE.

```
XX Lowe CR, Murray JAH, Squirrel DJ, White PJ;
PI WPI; 1995-344619/44.
DR
XX
XX New heat stable luciferase mutants - related DNA, vectors and
PT transformed cells, useful as reporter genes, assay labels and for
PT determ of ATP
XX
XX Disclosure; Page 23-25; 40pp; English.
XX
XX A doubly mutated luciferase of Photinus pyralis has leucine at
CC amino acid position 215 and an amino acid other than glutamic
CC acid at position 354. The luciferase, obtd. by site-directed
CC mutagenesis of the luc gene, shows improved thermostability
CC compared to the wild-type.
XX
SQ Sequence 550 AA;
Query Match 56.4%; Score 1588.5; DB 16; Length 550;
Best Local Similarity 55.5%; Pred. No. 7e-144;
Matches 306; Conservative 103; Mismatches 136; Indels 5; Gaps 3;
QY 1 MEE-ENIRHGERPRDIVHPGAGQOOLYQSLYKFAFPE--AIIDAHTNEVISYAFETS 57
Db 1 medaknkkpapfypdgedtagelqhkamkryalvpqtaftdahievntacyfems 60
QY 58 CRLAVSIQYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDELTHLNI 117
Db 61 vrleamkryglnthrvvcsenslqfmpvlgalfgavavapandiynerellnsni 120
QY 118 SKPTMFSSKALPILRVQONLSPIKVVWIDSMVDINGECVSTFVARTDHTFDPLS 177
Db 121 sqptvvfvekglqklnvqkklpiqkiiimdstdyvgfsgmyctfvtshlppgfneid 180
QY 178 FTKDFDPLEKIALIMSSSGTGLPKGVVLSHRSTIRFVHSRDIYGRTRVPQTSILS 237
Db 181 fvpsfdrdktialimnssgstgipkvalphrtlcvrfswardpifgnqilpdtailsv 240
QY 238 VPFHAFGMFTTSLFYVVGKLVVMLKKEGALFLKTIQNYKIPTIVAPPVNFVFLAKSPL 297
Db 241 vpfhgfmgfttlgylcgfrvmyrfveelflrslqdykqsalivptlfsfakatl 300
QY 298 VDQYDLSLSEVATGAPLGDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlsnlheiasgapskeveavakrfhlpgirgyltetsailtpxgddkpg 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGEIGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417
Db 361 avgvkvpfveakvvdldtktlgvnrqgelcvrgpmimgsvynpneatnaldkdgwlhs 420
QY 418 GDIGYVDEGNFIVDRILKELIKYGVQVAPAELENLLOHPSIADAGVTGVPDEFQGL 477
Db 421 gdiaywdehffvdrilsiklykgyvapaetesilqbpnifdgaaglpdddagel 480
QY 478 PAACVLESGLTLEKEYQDFIAQVTPTKHLRGVWFVDSIPKPTGKTKRKLREIF- 536
Db 481 paavvlehgtmtkeivdyasqvttakkirgsvvfvdevpkaltkdarkkirelli 540
QY 537 -AQRAPKSKL 545
Db 541 kakkggkskl 550
RESULT 2
AAR98515
ID AAR98515 standard; Protein; 550 AA.
XX
AC AAR98515;
XX
DT 11-MAR-1997 (first entry)
XX
```

```
DE P. pyralis variant luciferase E270X.
XX Luciferase; firefly; ATP; heat stability; luciferin; light.
XX
XX Photinus pyralis.
OS
XX Key Location/Qualifiers
FH Misc-difference 270
FT /label= E270X
FT /note= "Mutated to improve heat stability"
XX
XX WO9622376-A1.
XX
XX 25-JUL-1996.
XX
XX 19-JAN-1996; 96WO-GB000099.
XX
XX 24-APR-1995; 95GB-0008301.
XX 20-JAN-1995; 95GB-0001172.
XX
XX (MINA ) UK SEC FOR DEFENCE.
XX
XX Lowe CR, Murray JA, Squirrel DJ, White PJ;
PI WPI; 1996-354533/35.
XX N-PSDB; AAT36492.
XX
XX Mutant luciferase(s) with greater heat stability than wild-types -
XX useful in luminescent reagents
XX
XX Claim 9; Page 20-24; 39pp; English.
XX
XX This sequence represents a variant luciferase from the firefly, P.
XX pyralis. In this sequence the amino acid at position 270 has
XX been mutated such that it is not Glu. The luciferase has a
XX lower Km for the substrate ATP than native luciferase. The
XX luciferase therefore has a greater heat stability than wild-type
XX luciferase. The luciferase may be used in an assay to measure ATP
XX with luciferin, whereby the amount of light is related to the amount
XX of ATP present.
XX
XX Sequence 550 AA;
SQ
Query Match 56.4%; Score 1585.5; DB 17; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.4e-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;
QY 1 MEE-ENIRHGERPRDIVHPGAGQOOLYQSLYKFAFPE--AIIDAHTNEVISYAFETS 57
Db 1 medaknkkpapfypdgedtagelqhkamkryalvpqtaftdahievntacyfems 60
QY 58 CRLAVSIQYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDELTHLNI 117
Db 61 vrleamkryglnthrvvcsenslqfmpvlgalfgavavapandiynerellnsni 120
QY 118 SKPTMFSSKALPILRVQONLSPIKVVWIDSMVDINGECVSTFVARTDHTFDPLS 177
Db 121 sqptvvfvekglqklnvqkklpiqkiiimdstdyvgfsgmyctfvtshlppgfneid 180
QY 178 FTKDFDPLEKIALIMSSSGTGLPKGVVLSHRSTIRFVHSRDIYGRTRVPQTSILS 237
Db 181 fvpsfdrdktialimnssgstgipkvalphrtlcvrfswardpifgnqilpdtailsv 240
QY 238 VPFHAFGMFTTSLFYVVGKLVVMLKKEGALFLKTIQNYKIPTIVAPPVNFVFLAKSPL 297
Db 241 vpfhgfmgfttlgylcgfrvmyrfveelflrslqdykqsalivptlfsfakatl 300
QY 298 VDQYDLSLSEVATGAPLGDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlsnlheiasgapskeveavakrfhlpgirgyltetsailtpxgddkpg 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGEIGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417
```

Db	361	avgvkvvpff	eaavkvdtdgktlgvnrqgeicvrgpmimsgyvnpeactnalldkgw	lws	420
Qy	418	GDIGYYDEDCGNFFIVDR	KLRELKLYKGYQVAPAELENLLQHP	SIADAGVTGVPDEF	GGOL 477
Db	421	gdlaywdehffivdr	lkslikykgyvapaelssillqhnpi	fdagvaglpdddagel	480
Qy	478	PAACVVLESGKTLTEKEVODFTAAQVTP	PKHLRGVWFVDSIPKGTGKLIRKELREIP-		536
Db	481	paavvvlehgkmtkeivdyvasgvt	takkirggvfvdevpklgtgldarkirell		540
Qy	537	-AORAPKSKL	545		
Db	541	kakkggkskl	550		
RESULT	3				
ID	AAR50011		standard; Protein; 550 AA.		
XX	AAR50011				
AC	AAR50011;				
XX					
DT	29-SEP-1994		(first entry)		
XX					
DE	Firefly luciferase reporter protein.				
XX					
KW	Multicistronic expression unit; bicistronic vector system;				
KW	recombinant protein production; reporter protein; luciferase;				
KW	firefly.				
XX					
OS	Photinus pyralis.				
XX					
PN	W09405785-A.				
XX					
PD	17-MAR-1994.				
XX					
PF	26-AUG-1993; 93WO-EP02294.				
XX					
PR	27-AUG-1992; 92DE-4228458.				
XX					
PA	(BETIE) BEIERSDORF AG.				
PA	(GBFB) GBF GES BIOTECH FORSCHUNG GMBH.				
XX					
PI	Achterberg V, Dirks W, Dorschner A, Eichner W, Hauser H;				
PI	Meyer-Ingold W, Mielke H, Wirth M, Doerschner A;				
DR	WPI; 1994-101190/12.				
DR	N-PSDB; AAQ58732.				
XX					
PT	New multicistronic expression units - for producing equimolar				
PT	amts. of polypeptide(s) in mammalian cells as hosts				
XX					
PS	Example 1; Page 66-67; 109pp; German.				
XX					
CC	Reporter genes coding for secretory alkaline phosphatase (AAQ58731)				
CC	and for luciferase (AAQ38732) can be co-expressed using a bicistronic				
CC	system in which a sequence responsible for internal translation				
CC	start (designated 'IRES') is located between cistrons coding for				
CC	the two proteins.				
XX					
SQ	Sequence 550 AA;				
Query Match	56.3%;	Score	1584.5;	DB 15;	Length 550;
Best Local Similarity	55.5%;	Pred. No.	1.7e-143;		
Matches	305;	Conservative	103;	Mismatches	137;
				Indels	5;
				Gaps	
QY	1	MEE-ENIRHGERDRIHPG	SAGQQLYSLYKFAPEP--AIIDAHTNEVISYAQIFETS	57	
Db	1	medaknikgpapfyledgtageqlhkamkryalvpgtiaftdahleivnityaefems	60		
QY	58	CRUAVSIEQVGLNENN	VGVCSNNINFTNPVLAALYLGIPVATSNMDYTDGELTGH	117	

Db	61	vrlaeamkrvgltntshrivcsenslqffmpvlgalfigvavapandiynereelnsmai	121
Qy	118	SKPTIMFSKKALPLILRVOONLSPIKVVVVIDSMYDINGVECVSTFFVARYTDHTDPLS	177
Db	121	sqptvvfsvkqqlklnvqkkipqkiiimdsctdyqfgsmylvfvlshlppgfneyd	180
Qy	178	FTPKDFPDLKIALIMSSSTTGLPKGVVLSHRSLTIRFVHSRDPICYGTRTVDPOTSLSL	237
Db	181	fvpesfdrktialimassgstglpkgvalphrtacvrfshardpifgnqilpdtailsv	240
Qy	238	VPEHAFGMMFTTISYFVVGUKVVMVKKFEKALFKLTIONQYKPTIVVAPPVMVFLAKSPL	297
Db	241	vpfhgfgmfttlylgicgfrvvlmyrfeeeelfrsiqdykkgalsallvptlfsfakstl	300
Qy	298	VDQYDLSLSEVATGGAPLKGDAEAVAKRLKLPGLGIGYGLTETCCAVMITPHNAVKTG	357
Db	301	ldkydlsnlheiasggaplekevegeavakrfhlpgirqvgylcetttsallitpegddkpg	360
Qy	358	STGEPLPIYAKVLDNATGALGPGERGEICFQSEMITMKGYYNNPEATIDTIDKDGWLS	417
Db	361	avgkvvpffeaqvvdldgtklvgndrgelcvrpgmlmsgyvnonpeatnalldkdgwls	420
Qy	418	GDIGYDDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPSTADAGVTGVPDEFGGOL	477
Db	421	gdlaywdehffivdrklklykgyqvapaelesillqhnifdagvaglpdddagel	480
Qy	478	PAACVVLSESGKLTETEKEVODFIAAQVTPPTKHLRGVVVFVDSIPKPTGKLIIRKELRIF	536
Db	481	paavvvlehgkmtkeisvdyvasqvttaakklrgvvfvdevpkglgklidar:kireili	540
Qy	537	-AQRAPSKSL 545	
Db	541	kakkggkskl 550	
RESULT 4			
AAy08523			
ID	AAy08523 standard; Protein: 550 AA.		
XX	AAy08523;		
XX			
DT	03-AUG-1999 (first entry)		
XX	P. pyralis pTetLucI luciferase protein.		
DE			
KW	Firefly; luciferase; tetracycline; transcriptional control; TetR; Tet		
KW	tetracycline repressor; tetracycline promoter; luminescence; luxCDABE		
KW	insect; Tn10; medicine; dosage; cheese production; antibiotic; foodsa		
KW	allergy.		
XX			
OS	Photinus pyralis.		
XX			
PN	WO9925866-A1.		
XX			
PD	27-MAY-1999.		
XX			
PF	11-NOV-1998; 98WO-FI00873.		
XX			
PR	14-NOV-1997; 97FI-0004235.		
XX			
PA	(KARP/) KARP M.		
PA	(KORP/) KORPELA M.		
PA	(KURI/) KURITTU J.		
XX			
PI	Karp M, Korpela M, Kurittu J;		
XX			
DR	WPI, 1998-338015/28.		
DR	N-PSDE; AAV72416.		
XX			
XX			
PT	Assaying for tetracycline using recombinant prokaryotic cells		
XX			
PS	Example 2; Page 25-26; 67pp; English.		
XX			

Assaying for tetracycline using recombinant prokaryotic cells
Example 2; Page 25-26; 67pp; English.

CC This invention describes a novel tetracycline assay that uses recombinant
CC prokaryotic cells comprising a luciferase gene under the transcriptional
CC control of a tetracycline repressor and tetracycline promoter and
CC involves the detection of luminescence emitted from the cells. The assay
CC can be used to distinguish tetracycline from other microbial agents. The
CC invention also describes a novel plasmid comprising either the luxCDABE
CC genes, a tetracycline repressor (tetR) and a tetracycline promoter (tetA)
CC from Tn10, or the insect luciferase gene, a tetracycline repressor (tetR)
CC and a tetracycline promoter (tetA) from Tn10. The tetracycline assay
CC method can be used for the determination of tetracycline in a sample,
CC e.g. to study the dosage and penetration of the medicine. The method can
CC also be used to test cheese production, as cheese making bacteria are not
CC able to work in the presence of tetracycline. The method can also be used
CC to determine the presence or concentration of antibiotics in foodstuffs,
CC e.g. for allergic people. The present assay method does not rely on the
CC growth of microbes as do conventional tests, and so is much more rapid.
CC The present assay is also more sensitive, as even a small amount of
CC luminescence can be detected.

XX Sequence 550 AA;

Query Match 56.3%; Score 1584.5; DB 20: Length 550;
Best Local Similarity 55.5%; Pred. No. 1.7e-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOLYQSLYKFAFPE--AIIDAHTNEVTSYAQIFETS 57
Db 1 medaknkkgpapfypledgtagedqhkamkryalvpqtiatftdahlevnityaeefms 60
Qy 58 CRLAVSIQYGLNENNVGVCSENNINFFNPVLAALYIGIPVATSNMDYTGGELTGHNL 117
Db 61 vrlaeamkryglnthnrvvcenslqfmpvlgalfgavapandynereilnsmml 120
Qy 118 SKPTMFSSKKALPILRLVQOQLSFIKKVVDSDYINGVECVSTFVARYTDHTFDPLS 177
Db 121 sqptvfvskkgqlklnvqkklplqkllmdsktdyggfsgmyftvtshlppgfneyd 180
Qy 178 FTKPDFDPLEKIALMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpesfdrdktalimnsagtgipkqvalphrtacvrfshardpifgnqilpdtailsv 240
Qy 238 VPFHAFGCMFTLSYFVVGKLVVMLKKPEGALFLKTIQNYKIPTIVAPPVAVFLAKSPL 297
Db 241 vpfhbgfmgftlglcgrfvrvmrfeeeelfrlsqldqyqlsallvptlfsfakstl 300
Qy 298 VDQYDLSLITEVATGAGPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 ldkydlslitelasgagplskvegeavakrfhipgfrggygltetstailtpedgdkpg 360
Qy 358 STGRPLPIKARVLNATGKALPGGERGEICFQSSMIMKGYNNPEATIDTDRDGLWLS 417
Db 361 avgvkvpfekavvldtqtlgnqrgelevrgmngsgynnpaetnallidkgwlhs 420
Qy 418 GDIGYVDEGNFFIVDRUKELIKYGYQVAPAELENLLQHPSTADAGVTGVDPDFGQOL 477
Db 421 gdlaywdehffivdriskelikygyqvapaelesillqhnifdagvaglpddadagel 480
Qy 478 PAACVVLSSGKLTBEKQVDFTAAQVTPTKHLRGVGVFVDSIPKQPTCKLRKELREJF- 536
Db 481 paavvvvlgkmtkekeivdyvasqvtakkirgsvvfvdevpvgltgldarkireill 540
Qy 537 -AQRAPSKL 545
Db 541 kakkggakl 550

RESULT 5

AAB31172
ID AAB31172 standard; Protein; 550 AA.
XX
AC AAB31172;
XX

DT 02-APR-2001 (first entry)
XX Amino acid sequence of a luciferase protein.
DE
XX
KW Growth rate; death rate; reporter gene; luminescent protein;
KW fluorescent product; luciferase; green fluorescent protein; GFP.
OS Photinus pyralis.
XX
PN WO200075367-A1.
XX
XX 14-DEC-2000.
PD
XX 07-JUN-2000; 2000WO-FI00507.
PP
XX 07-JUN-1999; 99FI-0001296.
PR
XX (LILI/) LILIUS E.
PA (VIRT/) VIRT M.
PA
XX Lilius E, Virta M;
PI
XX WPI: 2001-061737/07.
DR N-PSDB; AAC86954.
DR
XX Assessing growth and death rates of a micro-organism in a desired
PT environment, by introducing 2 reporter genes encoding luminescent and
PT fluorescent products and detecting luminescent fluorescence -
XX Disclosure; Page 27-28; 32pp; English.
PS
XX The specification describes a method for assessing the growth rate
CC and death rate of a micro-organism within a predetermined time period
CC in a desired environment. The method comprises introducing at least
CC two reporter genes encoding luminescent and/or fluorescent products
CC into the micro-organisms, incubating the micro-organism within the
CC desired environment, and detecting luminescence and/or fluorescence
CC after a predetermined time period. Use of two different markers within
CC a micro-organism enables the differentiation between growth and death
CC rates. The method is used to assess the growth rate and death rate of
CC a micro-organism within a predetermined time period in a desired
CC environment. The present sequence represents a luciferase protein,
CC and is encoded by a plasmid which encodes luminescent and fluorescent
CC proteins, and is used in the method of the invention.
XX
SQ Sequence 550 AA;

Query Match 56.3%; Score 1584.5; DB 22: Length 550;
Best Local Similarity 55.5%; Pred. No. 1.7e-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOLYQSLYKFAFPE--AIIDAHTNEVTSYAQIFETS 57
Db 1 medaknkkgpapfypledgtagedqhkamkryalvpqtiatftdahlevnityaeefms 60
Qy 58 CRLAVSIQYGLNENNVGVCSENNINFFNPVLAALYIGIPVATSNMDYTGGELTGHNL 117
Db 61 vrlaeamkryglnthnrvvcenslqfmpvlgalfgavapandynereilnsmml 120
Qy 118 SKPTMFSSKKALPILRLVQOQLSFIKKVVDSDYINGVECVSTFVARYTDHTFDPLS 177
Db 121 sqptvfvskkgqlklnvqkklplqkllmdsktdyggfsgmyftvtshlppgfneyd 180
Qy 178 FTKPDFDPLEKIALMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpesfdrdktalimnsagtgipkqvalphrtacvrfshardpifgnqilpdtailsv 240
Qy 238 VPFHAFGCMFTLSYFVVGKLVVMLKKPEGALFLKTIQNYKIPTIVAPPVAVFLAKSPL 297
Db 241 vpfhbgfmgftlglcgrfvrvmrfeeeelfrlsqldqyqlsallvptlfsfakstl 300
Qy 298 VDQYDLSLITEVATGAGPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357

Db	301	idkydlsn,theiasgap,lskevgaeavkrflhplgrggyltetsalltpegdckpg	360
Qy	358	STGRPLPYIKAKVLONATKALCPGGERGICFCQSEMINKYYNNPEATIDTIDKDCWLHS	417
Ddb	361	avgkvvpfeakvddldgkltlgvqrgeclvrgpmnsyyvnnpeatnalldkdgvllhs	420
Qy	418	GDIGYYDEGNFTVDRLKELIKYGYVAPAELENLLQHPSIDADAGVTGVDPDEFGGOL	477
Ddb	421	galaywdehffivdrllksllykgyqvapaelesillqhnifdagvagipdddagel	480
Qy	478	PAACVYVLESQTKLTKEVQDFTAAQVTPPKHLRGGVVFVDSIPKGPOTKGLIRKELREIF-	536
Ddb	481	paavvvlehgkmtexelvyvasqvtacklrggvvfdevpkgltgkldarkireill	540
Qy	537	-AQRAPKSL 545	
Ddb	541	kakkggkskl 550	
RESULT	6		
AAAR83006			
ID	AAAR83006	standard; Protein; 550 AA.	
AC	AAAR83006;		
DT	12-FEB-1996	(first entry)	
DE	Firefly luciferase E354	engineered mutant.	
KW	Luciferase; thermostability; enzyme stabilization; firefly;		
KW	Photinus pyralis; luc gene; site-directed mutagenesis; label;		
KW	ATP analysis.		
XX	Synthetic.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 354		
FT	/note=	"amino acid at position 354 is any amino acid other than Glu, pref. also not Gly, pro or Asp, and is pref. Trp, Val, Leu, Ile, Asn or esp. Arg or Lys, their analogues or modified forms"	
XX	WO9525798-A1.		
XX	PN		
XX	PD		
XX	PP		
XX	22-MAR-1995;	95WO-GB00629.	
XX	20-JAN-1995;	95GB-0001170.	
XX	23-MAR-1994;	94GB-0005750.	
XX	(MINA) UK SEC FOR DEFENCE.		
PA			
XX	Low CR, Murray JAH, Squirrel DJ, White PJ;		
PI			
DR	WPI: 1995-344619/44.		
DR	N-PSDB; AAT05621.		
PT	New heat stable luciferase mutants - related DNA, vectors and		
PT	transformed cells, useful as reporter genes, assay labels and for		
PT	determin of ATP		
PS	Claim 7; Page 20-22; 40pp; English.		
XX			
CC	Novel luciferases are created by site-directed mutagenesis of the		
CC	P. pyralis luc gene (see AAT05621) to create substitutions at		
CC	amino acid position 354 (glutamate in the wild-type) of the encoded		
XX	protein (AAAR83006) that result in improved enzyme thermostability.		
XX	Sequence	550 AA;	

Query Match	56.3%	Score 1583.5	DB 16	Length 550
Best Local Similarity	55.5%	Pred. No. 2.1e-143		
Matches	305	Conservative 103	Mismatches 137	Indels 5
Gaps				
QY	1	MEE-ENRIGERPRDILVHPCSGAOLYQSILYKFASPE--AIIDAHTNEVISAQIFETS	57	
Db	1	medaknuikgppafpyldedgtageqihkamkryalvpgtiafdanheivniyaeefems	60	
QY	58	CRUAVSEIOYGLMENNVGVCSENNINFPNVLAAALYGLGIPVATSDMYTDGELTGHLL	117	
Db	61	vrlaeamkrylntnhrievcsensiqfempvlgalfigavapandiynereellnsmni	120	
QY	118	SKTFIMFSKKALPLILRVOQNLSTFKVWVIDSMTDINGECVSTFVARYTDHFDPLS	177	
Db	121	sqtvvfvfakgqiklnvgkklpiiqkiiimdsktdyqgfgmyftvtshlppgfneyd	180	
QY	178	FTPKDFDPLEKIALIMSSSTGTLPGVVLSHRSLTIRFVHSRDPYIGRTVPOTSLSL	237	
Db	181	fvpesfordktialimssgstgplkgvalphrtacvrfshardpifgqiipdtailav	240	
QY	238	VPEHHAFGMFTTILSYFVGLKVVMKLKFCGALFKLTIONYKIPTIVVAPPVVMFLAKSPL	297	
Db	241	vpthbgfmfttgiylogfrvvlmyrfeeflrsiqdykqsallvptlfsfakstl	300	
QY	298	VDQYDLSLSEVATGAPLGKDVAAVAKRLKLPGLIOGYGLTETCCAYMITPHNAVKTG	357	
Db	301	ldkydlnshelastgapiskevgaeavakrfhlpglrqgyigtettsailtctxgdkpg	360	
QY	358	STGRPLPYTKAKVLDNATGKALQPGSRGEICFQSEIMKGYNNPNPEATIDTIDKQGLHS	417	
Db	361	avkvvpvfeakvvdldtktlgyngrelcvrgpmimgsyvnppeatnalldkdgwlhs	420	
QY	418	GDYGYDEDNFTIVDRLEKIKYQYVAPAELENLLQHPISIADAGVTGVPDEFEGGOL	477	
Db	421	gdlaywdehdfivdrklsllkyqyvpaelesillghpnifdagvagipdddagel	480	
QY	478	PAACVYLESKTTTEKEVODFTAAOVTPTKHLRGVGVFVDSITPKGPTGKLIRKELREIF-	536	
Db	481	paavvlehgkmttekeivdyvasqvttakkirgvgvfvdevpkipgtgklidarkirelli	540	
QY	537	-AQRAPKSKL	545	
Db	541	kakggkkskl	550	
RESULT	7			
ID	AAR98518			
XX	AAR98518	standard; Protein; 550	AA.	
XX	AC			
XX	AC	AAR98518;		
XX	DT	11-MAR-1997	(first entry)	
XX	DE			
XX	DE	P. pyralis variant luciferase	E270K/E354K.	
XX	KW	Luciferase; firefly; ATP; heat stability; luciferin; light.		
XX	OS	Photinus pyralis.		
XX	XX	Key	Location/Qualifiers	
FX	FX	Misc-difference	270	
FT	FT	/label=	E270K	
FT	FT	Misc-difference	354	
FT	FT	/label=	E354K	
XX	XX	WO9622376-A1.		
PN	PN			
PD	PD	25-JUL-1996.		
XX	XX	19-JAN-1996;	96WO-CB000099.	
XX	XX	24-APR-1995;	95GB-0008301.	

Db 303 idkydlnlheliasggaplskevgeavakrfhlpgirggygttetsailitpkpgdfkpg 362

Qy 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSWIMKGYNNPEATIDTIDKDGWLHS 417

Db 363 avgvvpffekavvldtqktlgvnrqgelcvcrgpmlmsgyvnnpeatnalldkgwlhs 422

Qy 418 GDIGYDEDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPSTADAGVTGVPDEFGGOL 477

Db 423 gdiaywdehffivdrilkslikygyvapaeelesillqhpnlfdagvaglpddadagel 482

Qy 478 PAACVVLSESGKTLTEKEVODFTAAQVTPTKHLRGVGFVDSIPKGPCTGKLIRKELREIF- 536

Db 483 paavvvlehgtmtkekeivdyvasqvtakkirggyvfvdevpklgtgldarkirelli 542

Qy 537 -AORAPKSKL 545

Db 543 kakkggkskl 552

RESULT 9

AAR28127

ID AAR28127 standard; Protein; 815 AA.

XX AC AAR28127;

XX DT 12-MAR-1993 (first entry)

XX DE Lux: npt-II fusion protein.

XX KW Neomycin phosphotransferase-II; genetic selection; genetic marker;

XX KW gene recovery.

XX OS Synthetic.

XX PN WO9217593-A.

XX PD 15-OCT-1992.

XX PF 30-MAR-1992; 92WO-CA00139.

XX PR 28-MAR-1991; 91US-0676432.

XX PA (CANADA) NAT RES COUNCIL CANADA.

XX PI Crosby WL, Datla RSS, Hammerlindl JK, Selvaraj G;

XX WPI; 1992-366262/44.

XX DR N-PSDB; AAQ30000.

XX PT Fused gene with characteristics of component parts - comprising a

XX PT gene conferring conditional growth advantage and a marker gene,

XX PT used as a probe to select and isolate genetic elements

XX PS Disclosure; Page 33; 53pp; English.

XX CC The sequence is that of the fusion protein encoded by the firefly

XX CC (P.pyralis) luciferase:neomycin phosphotransferase-II chimeric gene.

XX CC See also AAR28125.

XX SQ Sequence 815 AA;

Query Match 56.2%; Score 1581.5; DB 13; Length 815;

Best Local Similarity 55.5%; Pred. No. 6.2e-143;

Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MPE-ENIRHGERPRDIVPGSGAQQLYQSLYKFASFPE--AIIDAHTNEVISAQIFETS 57

Db 1 medaknikkpgpfpyledgtagedlghkamkryalvpgtftadahlvniityaefems 60

Qy 58 CRLAVSIEQYGLNENNVGVCSNNINFPNVLALYLGIPVATSDMYTDGELTGLNLI 117

Db 61 vrlaamkrygintohrivvcsenslqfimpvlgalfgavavapandinyerellysmni 120

Qy 118 SKPTIMFSKKALPILIRVQONLSFIKKVWVIDSWYDINGECVSTFVARYTIDHTFDPLS 177

Db 121 sgptvfvfskkgqlilnvqkklplliqkllimdsktldyggfdgsmtyftvtshlppdfneyd 180

Qy 178 FTPKDFDPLEKIALIMSSSGTTGLPKGVVLSHRSLTIRFVHSRDPITYGTRTPVQTSILSL 237

Db 181 fypesfdrdktallmsssgstgplkqvalphrtacvrfshardplfgnqilpdtallsv 240

Qy 238 VPFHAFGMFTLTSFYVVLGVVLMKLFEGALFLKTIQNIKIPTIIVAPPVPMVFLAKSPL 297

Db 241 vpfhbgfmgmttlglylcgfrvvlmyrfeeeelfrlsgdykqlsqallvptlfsffakstl 300

Qy 298 VDQYDLSSLTEVATCGAPLGRDVAEAVAKRLKPLGIIQGYGLTETCCAVMIITPHNAVKTG 357

Db 301 idkydlnlheliasggaplskevgeavakrfhlpgirggygttetsailitpegddkpg 360

Qy 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSWIMKGYNNPEATIDTIDKDGWLHS 417

Db 361 avgvvpffekavvldtqktlgvnrqgelcvcrgpmlmsgyvnnpeatnalldkgwlhs 420

Qy 418 GDIGYDEDEGNFFIVDRKLKELIKYGYVAPAELENLLQHPSTADAGVTGVPDEFGGOL 477

Db 421 gdiaywdehffivdrilkslikygyvapaeelesillqhpnlfdagvaglpddadagel 480

Qy 478 PAACVVLSESGKTLTEKEVODFTAAQVTPTKHLRGVGFVDSIPKGPCTGKLIRKELREIF- 536

Db 481 paavvvlehgtmtkekeivdyvasqvtakkirggyvfvdevpklgtgldarkirelli 540

Qy 537 -AORAPKSKL 545

Db 541 kakkggkskl 550

RESULT 10

AAR54667

ID AAR54667 standard; Protein; 743 AA.

XX AC AAR54667;

XX DT 21-DEC-1994 (first entry)

XX DE Hepatitis C Virus core protein-luciferase fusion protein.

XX KW Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;

XX KW antisense; therapy; inhibition; viral protein precursor;

XX KW recombinant vaccinia virus; HCV core protein gene; firefly;

XX KW luciferase reporter gene; fusion construct.

XX OS Chimeric Hepatitis C virus.

XX OS Chimeric Photinus pyralis.

XX PN CA2104649-A.

XX PD 26-FEB-1994.

XX PF 23-AUG-1993; 93CA-2104649.

XX PR 25-AUG-1992; 92JP-0248796.

XX PR 03-MAR-1993; 93JP-0042736.

XX PA (SEKI/) SEKI M.

XX PI Honda Y, Seki M, Yamada E;

XX WPI; 1994-151836/19.

XX DR N-PSDB; AAQ65322.

XX PT Anti:sense oligo:nucleotide(s) complementary to the hepatitis C

XX PT virus genome - are useful as antiviral agents

XX Example 5; Page 244-250; 262pp; English.

CC A recombinant vaccinia virus which codes for a HCV core protein
CC fused to the firefly luciferase enzyme was constructed from PCR
CC amplified fragments. The construct (which encodes the fusion
CC protein sequence AAR54867) was useful for assaying the inhibitory
CC activity of various antisense oligonucleotides on HCV gene
CC translation; detection of luciferase activity provided a measure
CC of expression of the HCV core protein.
XX
SQ Sequence 743 AA;

Query Match 56.1%; Score 1577.5; DB 15; Length 743;
Best Local Similarity 55.3%; Pred. No. 1.3e-142;
Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYOSLYKFAFPE--AIIAHTNEVISYAOIFETS 57
DB 194 medaknikkppapfpledgtagedqlhamkryalvpgtlaftdchvievlyaeefems 253
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSDMYTDCGLTGHLMNI 117
DB 254 vrlaamkrygintnhrivvcsenslqfmpvlgalfgavapandlynereellnsmni 313
QY 118 SKPTTFSSKKALPALILRVQNLSEFIKKVVDSDMYDINGVCEVSTFVARYTDHTFDPLS 177
DB 314 sqptvfvskkgqlknvqkklpiqlkmdsktdyggfsgmyftvtshlppgfneyd 373
QY 178 FTPKDFDPLEKIALTMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRVPTQSILSL 237
DB 374 fypesfdrdktiallmnsstgtpkgyvalphrtacvrfshardpifgnqilpdcallsv 433
QY 238 VPFHAFGMFTTSLYFVVGKLVMLKKEGALFLKTIQNYKIPTIVVAPPVNVFLAKSPL 297
DB 434 vpfhbgfmfttglyicgrfvvmyrfeeeelfrslqdyklsallvptlfsffakstl 493
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMTPHNAVKTG 357
DB 494 ldkydlslheliasgagplskevgavakrfhlpgirggygttetsallitpegddkpg 553
QY 358 STGRPLPYIKARVLNATCKALGPGEIGEICQSEIMKMGYYNNPEATIDTIDKDWLHS 417
DB 554 avgvvpfpeakvvdltgktlgnvqrgelevrgmmsygnnpeatnallidkgwlhs 613
QY 418 GDIGYDEDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPSTADAGVTGVPDEFGQOL 477
DB 614 gdlaywdehdfvdlklslikygyqvapaelesillqhnifdagvagipddadagel 673
QY 478 PAACVVLSESGKLTETKEVQDFIAAOVTPTKHLRGVGVFVDSIPKGTGKLRKELREIF- 536
DB 674 paavvvhgktmtekeivdyvasqvtacklrggvvfvdevpklgtkldarkirell 733
QY 537 -AQRAPSKL 545
DB 734 kakkggkkl 743

RESULT 11
AAR72801
ID AAR72801 standard; Protein; 743 AA.

XX AAR72801;
XX AAR72801;
DT 04-DEC-1995 (first entry)
DE Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.
XX Vaccinia virus vector; firefly luciferase; hepatitis C virus; T7N1-19;
KW antiviral agent; poliovirus; human rhinovirus;
KW internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;
KW foot and mouth disease.
XX Synthetic.
XX

PN JP07069899-A.
XX 14-MAR-1995.
XX 02-SEP-1993; 93JP-0241973.
XX 02-SEP-1993; 93JP-0241973.
XX (MITU) MITSUBISHI KASEI CORP.
XX WPI; 1995-144713/19.
DR N-PSDB; AAR86799.
XX Antiviral agent comprising component which disrupts viral gene
PT translation - used for the selective inhibition of e.g. Hepatitis
PT C virus, polio:virus and human rhinovirus
XX Disclosure; Pages 19-22; 23pp; Japanese.
XX AAR86788 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which
CC disrupts viral gene translation, by preventing the binding of
CC the viral mRNA to the internal ribozyme entry site. It was
CC used in the construction of an antiviral agent detecting vector
CC (AAR86799 which encodes AAR72801), which comprises a vaccinia virus
CC vector, a vector containing T7N1-19 and a firefly luciferase gene.
CC The antiviral agent can be used to treat HCV, poliovirus, cerebral
CC cardio-hepatitis, human rhinovirus and foot and mouth disease viral
CC infections.
SQ Sequence 743 AA;

Query Match 56.1%; Score 1577.5; DB 16; Length 743;
Best Local Similarity 55.3%; Pred. No. 1.3e-142;
Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYOSLYKFAFPE--AIIAHTNEVISYAOIFETS 57
DB 194 medaknikkppapfpledgtagedqlhamkryalvpgtlaftdchvievlyaeefems 253
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSDMYTDCGLTGHLMNI 117
DB 254 vrlaamkrygintnhrivvcsenslqfmpvlgalfgavapandlynereellnsmni 313
QY 118 SKPTTFSSKKALPALILRVQNLSEFIKKVVDSDMYDINGVCEVSTFVARYTDHTFDPLS 177
DB 314 sqptvfvskkgqlknvqkklpiqlkmdsktdyggfsgmyftvtshlppgfneyd 373
QY 178 FTPKDFDPLEKIALTMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRVPTQSILSL 237
DB 374 fypesfdrdktiallmnsstgtpkgyvalphrtacvrfshardpifgnqilpdcallsv 433
QY 238 VPFHAFGMFTTSLYFVVGKLVMLKKEGALFLKTIQNYKIPTIVVAPPVNVFLAKSPL 297
DB 434 vpfhbgfmfttglyicgrfvvmyrfeeeelfrslqdyklsallvptlfsffakstl 493
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYGLTETCCAVMTPHNAVKTG 357
DB 494 ldkydlslheliasgagplskevgavakrfhlpgirggygttetsallitpegddkpg 553
QY 358 STGRPLPYIKARVLNATCKALGPGEIGEICQSEIMKMGYYNNPEATIDTIDKDWLHS 417
DB 554 avgvvpfpeakvvdltgktlgnvqrgelevrgmmsygnnpeatnallidkgwlhs 613
QY 418 GDIGYDEDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPSTADAGVTGVPDEFGQOL 477
DB 614 gdlaywdehdfvdlklslikygyqvapaelesillqhnifdagvagipddadagel 673
QY 478 PAACVVLSESGKLTETKEVQDFIAAOVTPTKHLRGVGVFVDSIPKGTGKLRKELREIF- 536
DB 674 paavvvhgktmtekeivdyvasqvtacklrggvvfvdevpklgtkldarkirell 733
QY 537 -AQRAPSKL 545


```
XX CC The present sequence is encoded by the modified firefly luciferase
CC cDNA luc+, which can be used as a reporter gene in experimental
CC biological systems, e.g. cell cultures, transgenic plants and
CC animals, and in cell free expression systems. Removal of the cDNA's
CC peroxisomal translocation sequence yields a cytoplasmic form of
CC the enzyme with optimal substrate availability. Potentially
CC interfering restriction sites and genetic regulatory sites are
CC removed, and codon usage for mammalian cells is improved. Absence
CC of potential N-glycosylation sites minimises post-translational
CC modification.
XX SQ Sequence 550 AA;

Query Match 55.8%; Score 1568.5; DB 18; Length 550;
Best Local Similarity 54.7%; Pred. No. 5.9e-142;
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRHGERPRDIVHPGAGQOOLYQSLYKFAFPPE--AIIDAHTEVVISYAOIFETS 57
Db 1 medaknikkppapfpledgtageqlhkamkryalvpgtiaftdahlevdityaeyfems 60

Qy 58 CRLAVSIEQYGLNENNVGVCSENNIFNPNVLAALYLGIPVATSNMYTDGELTGHLMNI 117
Db 61 vrleamkrygintnhrivvcsenslqfmpvlgalfgavavapandiynerellnsmgi 120

Qy 118 SKPTIMFSKXKALPILRLVQOQLSPFKKVVVIDSMDYINGVECVSTFVARVTDHRTDPLS 177
Db 121 sqptvvfsvsk9qlklnvqkklpliqkllimdshtdyggfsgmyfvtshlppgfneyd 180

Qy 178 FTKPDFLEXIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpsfdrdktalimnssgstglpkvalphrtacvrfshardpifgnqilpdtailsv 240

Qy 238 VPFHAFQCFMTTLSYFVGLKVVMLKKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Db 241 vpfhgfmgfntlgylcgfrvvlmyrfeeeelfrslqdyklsallvptlfsfakstul 300

Qy 298 VDOYDLSLSEVATGCGAPLGDVAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlnlhelasgagplskveavakrfhlpgirggyglctetsallitpegddkpg 360

Qy 358 STGRPLPYIKARVLNATGKALGPGERGEICFQSEIMKMGYNPNPEATIDTKDGLWLS 417
Db 361 avgvkvpfesakvvdldgktlgnvrgelcvgmimgynpnpeatnalidkgwlhs 420

Qy 418 GDIGYDDEGNEFFIVDRKLKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFEGQOL 477
Db 421 gdiaywdedehffivdrilksilkygyqvapaelesillqhnifdagvaglpddadagel 480

Qy 478 PAACVPLESGKTLTEKEVQDFTAAQVTPTKHLRGVGVFVDSIPKQPTCKLIRKELREIFA 537
Db 481 paaavvlehgktmtekeivdyvasqvttakkirgvgvfvdevpvgltgkldarkirelli 540

Qy 538 QRAPKSKL 545
Db 541 kakkggki 548

RESULT 14
ID AAB35055
XX AAB35055 standard; Protein: 550 AA.
XX AC AAB35055;
XX 27-MAR-2001 (first entry)
XX Plant signal transduction pathway reporter vector p2P-ALE protein.
XX Plant signal transduction pathway; reporter vector;
XX luciferase; herbicide resistance; pest resistance.
```

```
OS Synthetic.
XX WO200071668-A2.
XX 30-NOV-2000.
XX 22-MAY-2000; 2000WO-US14041.
XX 20-MAY-1999; 99US-0136145.
XX (UYNE-) UNIV NEBRASKA.
XX Rhoades DN;
XX WPI: 2001-025143/03.
XX N-PSDB; AAC56931.
XX Identifying genes encoding signal transduction components useful for
XX producing transgenic plants, by transforming plants with vector
XX encoding reporter gene, mutating the plant and identifying genes from
XX the mutant -
XX Disclosure; Page 47-49; 54pp; English.
XX The present invention provides a method for identifying genes encoding
XX components of plant signal transduction pathways between mitochondrial
XX function and metabolic status and nuclear gene expression. This involves
XX the transformation of a plant with a vector encoding a reporter gene
XX linked to an AOX promoter, mutagenesis of the plant to increase
XX expression of the reporter gene and then determining the identity of the
XX signal transduction gene of interest. This is useful in the production of
XX transgenic plants with increased productivity, herbicide, stress and pest
XX resistance.
XX Sequence 550 AA;

Query Match 55.8%; Score 1568.5; DB 22; Length 550;
Best Local Similarity 54.7%; Pred. No. 5.9e-142;
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRHGERPRDIVHPGAGQOOLYQSLYKFAFPPE--AIIDAHTEVVISYAOIFETS 57
Db 1 medaknikkppapfpledgtageqlhkamkryalvpgtiaftdahlevdityaeyfems 60

Qy 58 CRLAVSIEQYGLNENNVGVCSENNIFNPNVLAALYLGIPVATSNMYTDGELTGHLMNI 117
Db 61 vrleamkrygintnhrivvcsenslqfmpvlgalfgavavapandiynerellnsmgi 120

Qy 118 SKPTIMFSKXKALPILRLVQOQLSPFKKVVVIDSMDYINGVECVSTFVARVTDHRTDPLS 177
Db 121 sqptvvfsvsk9qlklnvqkklpliqkllimdshtdyggfsgmyfvtshlppgfneyd 180

Qy 178 FTKPDFLEXIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpsfdrdktalimnssgstglpkvalphrtacvrfshardpifgnqilpdtailsv 240

Qy 238 VPFHAFQCFMTTLSYFVGLKVVMLKKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Db 241 vpfhgfmgfntlgylcgfrvvlmyrfeeeelfrslqdyklsallvptlfsfakstul 300

Qy 298 VDOYDLSLSEVATGCGAPLGDVAEAVAKRLKLPGLIGQYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlnlhelasgagplskveavakrfhlpgirggyglctetsallitpegddkpg 360

Qy 358 STGRPLPYIKARVLNATGKALGPGERGEICFQSEIMKMGYNPNPEATIDTKDGLWLS 417
Db 361 avgvkvpfesakvvdldgktlgnvrgelcvgmimgynpnpeatnalidkgwlhs 420

Qy 418 GDIGYDDEGNEFFIVDRKLKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFEGQOL 477
Db 421 gdiaywdedehffivdrilksilkygyqvapaelesillqhnifdagvaglpddadagel 480
```

Qy	478	PAACVYLESGKTLTKEVODFTAAOVTPTKHLRGGVVFVDSITPKGPTGLRKELREIFA	537
Db	481	paavvlehgktmtekeivdyasqvttakkirggyvfvdevpkgitgkidakirrelli	540
Qy	538	QRAPKSKL	545
		:	:
Db	541	kakggki	548

RESULT 15
AAW32937
ID AAW32937 standard; Protein; 551 AA.

AC AAW32937;

DT 22-JAN-1998 (first entry)

DE Modified firefly luciferase luc+NF.

Firefly; luciferase; mutant; reporter gene; cell culture;
transgenic plant; transgenic animal; cell free expression system;
peroxisomal translocation sequence; cytoplasmic form; luc-NF;
restriction site; genetic regulatory site; N-glycosylation site;
post-translational modification.

OS Photinus pyralis.

XX
XX

FT Misc-difference 2..3

FT Misc-difference 51

Misc-difference 120

FT Misc-difference 549

FT Misc-difference 550

FT Misc-difference 551

XX
XX

XX

XX

XX

XX

[illegible]XX
XX

DR N-PSDB; AAT88244.

PT Modified lucifera

XX	Disclosure; Columns 31-36; 31pp; English.
PS	

CC The present sequence is encoded by the mo

biological systems, e.g. cell cultures, transgenic plants and animals, and in cell free expression systems. Removal of the peroxisomal translocation sequence yields a cytoplasmic form of the enzyme with optimal substrate availability. Potentially interfering restriction sites and genetic regulatory sites are removed, and codon usage for mammalian cells is improved. Absence of potential N-glycosylation sites minimises post-translational modification. The variation at 5'-end of the cDNA was designed for easier construction of amino-terminal fusions.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:00:31 ; Search time 14.58 Seconds
(without alignments)
913.028 Million cell updates/sec

Title: US-09-993-874-2
Perfect score: 2812

Sequence: 1 MESENTHGERPRDIHVPGS.....KLIRKELREIFAORAPKSKL 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1588.5	56.5	550	4	US-08-718-425-5
2	1588.5	56.5	550	4	US-08-875-277A-4
3	1585.5	56.4	550	4	US-08-875-277A-2
4	1584.5	56.3	550	3	US-08-867-352-23
5	1584.5	56.3	550	4	US-09-380-061B-21
6	1584.5	56.3	815	1	US-08-122-520C-9
7	1583.5	56.3	550	4	US-08-718-425-2
8	1580.5	56.2	549	1	US-08-354-240A-2
9	1576.5	56.1	550	4	US-09-380-061B-6
10	1568.5	55.8	550	1	US-08-354-240A-4
11	1568	55.8	550	1	US-08-354-240A-6
12	1542.5	54.9	561	2	US-08-474-169-8
13	1534	54.6	547	4	US-09-380-061B-20
14	1524.5	54.2	552	3	US-09-111-752-5
15	1521.5	54.1	552	3	US-09-111-752-7
16	1515	53.9	552	1	US-08-231-729B-6
17	1492.5	53.1	552	3	US-09-111-752-10
18	1491.5	53.0	548	1	US-07-675-211-2
19	1491.5	53.0	548	1	US-07-903-047-2
20	1491.5	53.0	548	1	US-08-076-042-2
21	1491.5	53.0	548	4	US-09-380-061B-14
22	1483	52.7	548	4	US-09-380-061B-18
23	1464.5	52.1	636	3	US-09-111-752-14
24	1464.5	52.1	636	2	US-08-460-934-9
25	1464.5	52.1	636	2	US-08-782-118-9
26	1463.5	52.0	548	2	US-08-460-934-2
27	1463.5	52.0	548	2	US-08-782-118-2

Sequence 6, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 16, Appli
Sequence 8, Appli
Sequence 1, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 11, Appli
Sequence 28, Appli
Sequence 29, Appli
Sequence 79, Appli
Sequence 97, Appli
Sequence 79, Appli
Sequence 97, Appli

ALIGNMENTS

RESULT 1
US-08-718-425-5
; Sequence 5, Application US/08718425
; Patent No. 6132983
; GENERAL INFORMATION:
; APPLICANT: Lowe, Christopher R.
; APPLICANT: White, Peter J.
; APPLICANT: Murray, James A. H.
; APPLICANT: Squirrell, David J.
; TITLE OF INVENTION: LUCIFERASES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6132983th Glebe Road, 8th Floor
; CITY: Arlington
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,425
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-539
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-718-425-5

Query Match 56.5%; Score 1588.5; DB 4; Length 550;
Best Local Similarity 55.6%; Pred. No. 5.4e-154;
Matches 306; Conservative 103; Mismatches 136; Indels 5; Gaps 3;

OY 1 MEE-ENIRHGERPRDIHVPGSAGQQLYQSLYKFASFPPE--AIIDAHTNEVISVAQIFETS 57

DB 1 MEDAKNIKKGAPPYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNTIYAEFFENS 60

Db	241	VPFHGEGMTTLLGSLGGRVVMYRFXEKLRLSDQYKIQSALLVPTLFSFAKSTL	300
Qy	298	VDOYDLSLSEVATGGAPLGDVAEAVAKRLKPGIIOGVGLTETCCAVMIAPHNAVKTG	357
Db	301	IDKIDLSNLHLEISGGAPLSKEVGEAVAKRPHLPGRIOGVGLTETTSAILITPEGDDKPG	360
Qy	358	STGRPLPYIKAKVLNDNATGKALGPGERGEICFOSEIMKGYNNPEATIDTIDKDGWLHS	417
Db	361	AVGWVPFEAKVVDLDTGKTLGVNQRGELCVRPMIMSGYNNPEATNALIDKDGWLHS	420
Qy	418	GDIGYDDEGNFIVDRBLKELIKYGYVAPAELEWLLLOHPSIADAGVTGVPDEFGGOL	477
Db	421	GDIAWDEDEHFFIVDRKLSIKYGYVAPAELESILLQHPNIFDAGVAGLPDDDAGEL	480
Qy	478	PAACVWLESGKTLTEKRVQDFIAAQVTPTKHLRGGVVFVDSIPKGP TGKILRKELREIF	536
Db	481	PAAVVLEHGKTFTEKRVQDFIAAQVTPTKHLRGGVVFVDSIPKGP TGKILRKELREIF	536
Qy	537	-AQRAPSKL	545
Db	541	KAKKGGSKL	550
RESULT	4		
US	-08-867-352-23		
;	Sequence 23, Application US/08867352		
;	Patent No. 6060273		
;	GENERAL INFORMATION:		
;	APPLICANT:		
;	TITLE OF INVENTION: Multicistronic expression units and their use		
;	NUMBER OF SEQUENCES: 25		
;	COMPUTER READABLE FORM:		
;	MEDIUM TYPE: Floppy disk		
;	COMPUTER: IBM PC compatible		
;	OPERATING SYSTEM: PC-DOS/MS-DOS		
;	SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)		
;	CURRENT APPLICATION DATA:		
;	APPLICATION NUMBER: US/08/867,352		
;	FILING DATE:		
;	CLASSIFICATION:		
;	PRIOR APPLICATION DATA:		
;	APPLICATION NUMBER: 08/387,847		
;	FILING DATE:		
;	INFORMATION FOR SEQ ID NO: 23:		
;	SEQUENCE CHARACTERISTICS:		
;	LENGTH: 550 amino acids		
;	TYPE: amino acid		
;	TOPOLOGY: linear		
;	MOLECULE TYPE: protein		
;	US-08-867-352-23		

Query Match	56.3%	Score 1584.5;	DB 3;	Length 550;
Best Local Similarity	55.5%	Pred. No. 14e-153;		
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps				
y y	1	MEE-ENIRHERPRDIVHPGSAGOOYLKYKFASPE--AIDAHTNEVSYAQIFETS	57	
b b	1	MEDAKNIKKGPAFFPLEDGAGEQLHKAMRYALVPCTIAFTADHAIEVNITVAEYFEMS	60	
y y	58	CRLAVSIEOYGLNENNVCVCSSENNFNFPDLAALVILGPVATSDMDYTDELGTCHLNI	117	
b b	61	VRLAEAMKRGUNTHRIIVCVSCNSLOQFMVLGALFIGVAVAPANDIYMERELLMSNI	120	
y y	118	SKPTIMFSKKLPILIRYQQNLSPIKRVVIDSGMYDINGVECVSTPVARVYTDHRTDPLS	177	
b b	121	SOPTVVVSVSKGLQILNVOKLPIQIIIMDSKTDYQGFSWYTPVTSHLPFGGFNEYD	180	
y y	178	FTEPKDFDPLEKIALIMSSGGTGTGPKGVVLGSHRSULTIRFVHSRDIYXGTRTVQTSLSL	237	
b b	181	FYPESDRDKTIALIMNSSGGTGPKGVALPHRTCAVFYSHARDPIFGNOIIPDTAISLV	240	

QY 238 VPFHAFGFTTSLSYFVGLKVVMLKKPEGALFLKTIQNYKIPTIVAPPVMMVFLAKSPL 297
Db 241 VPFHGFSGFTTGLYLCIGFRVLMYRFEELFLASLDYKIOSALLVPTLFFFAKSTL 300
QY 298 VDQYDLSLSEVATGAGPLGKDVAAVAKRLKPLGLIOGYGLTETCCAVMTPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGAPLSKEVGAARFHLPGIROGYGLTETTSAILITPEGDOKPG 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417
Db 361 AVGVVPPFEAKVLDGTGKLVNQRELVCVRGPMKSGYNNPEATNALIDKDWLHS 420
QY 418 GDIGYDEDEGFFIVDRKLKELIKYGYVAPAELENLLQHPSIADAGVTGVPDEFGGOL 477
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVQDFTAAQVTPTKHLRGGVWFVDSIPKGTGKLIRKELREIF- 536
Db 481 PAAVVVLEHGKTMTEKEIVDVYASQVTTAKKLKRGVWFVDEVPKGLTGKLDARKIREILI 540
QY 537 -AQRAPKSKL 545
Db 541 KAKGKSKL 550

RESULT 5
US-09-380-061B-21
; Sequence 21, Application US/09380061B
; Patent No. 6265177
; GENERAL INFORMATION:
; APPLICANT: SQUIRRELL, DAVID JAMES
; WHITE, PETER JOHN
; LOWE, CHRISTOPHER ROBIN
; MURRAY, JAMES AUGUSTUS HENRY
; TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,061B
; FILING DATE: 25-Aug-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01026
; FILING DATE: 7-APR-1998
; APPLICATION NUMBER: GB 9707468.8
; FILING DATE: 11-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B. J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 124-725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-380-061B-21

Query Match 56.3%; Score 1584.5; DB 4; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.4e-153;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;
QY 1 MBE-ENIRHGERPRDIVHPGSAGQOOLYOSLYKFPASPE--ATIDAHTNEVISYAOIFETS 57
Db 1 MEDAKNIKKPAPPEPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNTIYAEVEMS 60
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNPVLAALYIGIPVATSDMYTDCGLTGLHNI 117
Db 61 VRLAAMKRYGLNTHRIWVCSNLSQPFMPVGLFTGVAVAPANDYNERELLSMNI 120
QY 118 SKPTIMFSKSLPALILRVQONLSPIKVVVIDSMDYINGVCEVSTFVARYTDHTEFPLS 177
Db 121 SPTVVVFSKGLQKILNVQKLPILQKIIIMDSKTDVGGQSMYTFVTSHPGFEYD 180
QY 178 FTKPDFPLEKIALIMSSGTTGLPKGVVLSHRSITIRFVHSRDIYGTRTVPQTSILSL 237
Db 181 FVPSFDRDKTIALIMSSGSTGLPKGVVALPHRTACVRFSHARDPIFGNQIIPDTAILSV 240
QY 238 VPFHAFGFTTSLSYFVGLKVVMLKKPEGALFLKTIQNYKIPTIVAPPVMMVFLAKSPL 297
Db 241 VPFHGFSGFTTGLYLCIGFRVLMYRFEELFLASLDYKIOSALLVPTLFFFAKSTL 300
QY 298 VDQYDLSLSEVATGAGPLGKDVAAVAKRLKPLGLIOGYGLTETCCAVMTPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGAPLSKEVGAARFHLPGIROGYGLTETTSAILITPEGDOKPG 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417
Db 361 AVGVVPPFEAKVLDGTGKLVNQRELVCVRGPMKSGYNNPEATNALIDKDWLHS 420
QY 418 GDIGYDEDEGFFIVDRKLKELIKYGYVAPAELENLLQHPSIADAGVTGVPDEFGGOL 477
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVQDFTAAQVTPTKHLRGGVWFVDSIPKGTGKLIRKELREIF- 536
Db 481 PAAVVVLEHGKTMTEKEIVDVYASQVTTAKKLKRGVWFVDEVPKGLTGKLDARKIREILI 540
QY 537 -AQRAPKSKL 545
Db 541 KAKGKSKL 550

RESULT 6
US-08-122-520C-9
; Sequence 9, Application US/08122520C
; Patent No. 5639663
; GENERAL INFORMATION:
; APPLICANT: Crosby, William L.
; APPLICANT: Datla, Raju S.S.
; APPLICANT: Hammerlindl, Joseph K.
; APPLICANT: Selvaraj, Gopalan
; TITLE OF INVENTION: BIFUNCTIONAL GENETIC MARKERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. Wayne Anderson
; STREET: Montreal Rd.
; CITY: Ottawa
; STATE: ON
; COUNTRY: CANADA
; ZIP: K1A 0R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,520C
; FILING DATE: 11/29/92
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Anderson, J. Wayne
REGISTRATION NUMBER: 28,158
REFERENCE/DOCKET NUMBER: 10013-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 993-3899
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 815 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-122-520C-9

Query Match 56.3%; Score 1584.5; DB 1; Length 815;
Best Local Similarity 55.5%; Pred. No. 2.7e-153;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;
QY 1 MEE-ENIRGERPRDIVHPGSAQQLYQSLYKFASFPPE--AIIDAHTNEVSYAQIFETS 57
DB 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPOTIAFTDAHIEVNTIYAEYFEMS 60
QY 58 CRLAVSIEOYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSDMYTDCGLTGHNL 117
DB 61 VRLAEAMKRYGLNTHRIYVCSNSLQFPVGLGALFVGAVAPANDIYNERELNSMNI 120
QY 118 SKPTIMFSSKKALPLILRVQONLSFIKKVVIDSMYDINGVECVSTFVARYTDHFTDPLS 177
DB 121 SPTVVFVSKKGLQILNVOKKLPITQIIIMDSKTDYQGFOSMTFVTSHPGDFNEID 180
QY 178 FPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRTPQTSILSL 237
DB 181 FVPESFDRDKTIALIMNSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240
QY 238 VPFHAFGFTTSLYFVGLKVMKPKFEGALFLKTIQNYKIPTIVVAPVAVFLAKSPL 297
DB 241 VPFHGFPGNFTTGLYLCGRFVLMYRFEELFLSLQDYKIQSALLVPTLFSFFAKSTL 300
QY 298 VQOYDLSLSEVATGAPGKDVAAVAKRLKPLGIIQGYGLTETCCAVMITPHNAVKTG 357
DB 301 IDKYDLSNLHETASGAPLSKEVGEAVAKRFLPGTRQCYGLTETTSAILITPEGDGKPG 360
QY 358 STGRPLPYIKAKVLDNATKALGPGERGICFQSEIMKGYNNPEATIDTIDKQGLWLS 417
DB 361 AVGVVVFPEAKVVDLDTGTLGVNQRGELCVRGPMISGVYNNPEATNALIDKQGLWLS 420
QY 418 GDIGYDEGNFFIVDRKELIKYGYQVAPAELENLLQHPISADAGVTGVPDFFGGOL 477
DB 421 GDIAWDEDEHFFIVDRKLSLKYGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVODFTAAQVTPPKHLRGVGVFVDSIPKGTGKLIRKELREIF- 536
DB 481 PAAVVLSEHGKTMTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 540
QY 537 -AQRAPKSKL 545
DB 541 KAKKGKSKL 550

RESULT 7

US-08-718-425-2
Sequence 2, Application US/08718425
Patent No. 6132983
GENERAL INFORMATION:
APPLICANT: Love, Christopher R.
APPLICANT: White, Peter J.
APPLICANT: Murray, James A. H.
APPLICANT: Squirell, David J.
TITLE OF INVENTION: "LUCIFERASES"
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye

STREET: 1100 No. 6132983th Glebe Road, 8th Floor
CITY: Arlington
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718.425
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Arthur R. Crawford
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-539
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-718-425-2

Query Match 56.3%; Score 1583.5; DB 4; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.8e-153;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;
QY 1 MEE-ENIRGERPRDIVHPGSAQQLYQSLYKFASFPPE--AIIDAHTNEVSYAQIFETS 57
DB 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPOTIAFTDAHIEVNTIYAEYFEMS 60
QY 58 CRLAVSIEOYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSDMYTDCGLTGHNL 117
DB 61 VRLAEAMKRYGLNTHRIYVCSNSLQFPVGLGALFVGAVAPANDIYNERELNSMNI 120
QY 118 SKPTIMFSSKKALPLILRVQONLSFIKKVVIDSMYDINGVECVSTFVARYTDHFTDPLS 177
DB 121 SPTVVFVSKKGLQILNVOKKLPITQIIIMDSKTDYQGFOSMTFVTSHPGDFNEID 180
QY 178 FPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRTPQTSILSL 237
DB 181 FVPESFDRDKTIALIMNSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240
QY 238 VPFHAFGFTTSLYFVGLKVMKPKFEGALFLKTIQNYKIPTIVVAPVAVFLAKSPL 297
DB 241 VPFHGFPGNFTTGLYLCGRFVLMYRFEELFLSLQDYKIQSALLVPTLFSFFAKSTL 300
QY 298 VQOYDLSLSEVATGAPGKDVAAVAKRLKPLGIIQGYGLTETCCAVMITPHNAVKTG 357
DB 301 IDKYDLSNLHETASGAPLSKEVGEAVAKRFLPGTRQCYGLTETTSAILITPEGDGKPG 360
QY 358 STGRPLPYIKAKVLDNATKALGPGERGICFQSEIMKGYNNPEATIDTIDKQGLWLS 417
DB 361 AVGVVVFPEAKVVDLDTGTLGVNQRGELCVRGPMISGVYNNPEATNALIDKQGLWLS 420
QY 418 GDIGYDEGNFFIVDRKELIKYGYQVAPAELENLLQHPISADAGVTGVPDFFGGOL 477
DB 421 GDIAWDEDEHFFIVDRKLSLKYGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480
QY 478 PAACVVLSESGKTLTEKEVODFTAAQVTPPKHLRGVGVFVDSIPKGTGKLIRKELREIF- 536
DB 481 PAAVVLSEHGKTMTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 540
QY 537 -AQRAPKSKL 545
DB 541 KAKKGKSKL 550

Query Match 56.1%; Score 1576.5; DB 4; Length 550;
Best Local Similarity 55.3%; Pred. No. 9.2e-153;

Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGAGOOYQSLYKFASFPE--AIIDAHTNEVSYAQIFETS 57
Db 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDVITYAEYFEMS 60

Qy 58 CRLAVSIOYGLNENNVGVCSENNINFPNVLAAALYLGIPVATNSDMYTDELTHLNI 117
Db 61 VRLAEAMKRYGLNTHRVVCSNSLOQFMPVLGALFICVAVAPANDIYNRELLNSMI 120

Qy 118 SKPTIMESSKKALPLILRVQONLSFIKKVWIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Db 121 SPTVVFVSKKGLQKILNVOKKLPDIIQIIIMDSKTDYQFOSMTFTVTSHPGFNEYD 180

Qy 178 FTKPDFDPLEKIALIMSSGTTGPKGVLSHRSLTIRFVHSRDPYIGTRVTPOTSILSL 237
Db 181 FVPSFDRDKTIALIMSSGTTGPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240

Qy 238 VPFHAFGFMFTLSFVYVGLKVMKPEGALFKTIQNYKIPTIVAPPVNVFLAKSPL 297
Db 241 VPFHGFQFMFTLGLYICGFRVLMYREEEELFRLSLODYKIOSALLVPTLFSFAKSTL 300

Qy 298 VDQYDLSLSEVATGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRFLPCHIROGYGLTETTSAILITPEGDDKPG 360

Qy 418 GDIGYDEGDNFFIVDRKELIKYGYVAPAELENLLQHPNIFDAGVAGLPDDAGEL 477
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480

Qy 478 PAACVYLESGLTKEKEVQDFIAAQVTPKHLRGVGFVDSIPKGPCTGLKRLKELREIFA 536
Db 481 PAADVLEHGKMTKEIYDVASQVTTAKKLRGGVGFVDEVPKGLTKLDARKIREILI 540

Qy 537 -AORAPSKL 545
Db 541 KAKRGKSKL 550

RESULT 10
US-08-354-240A-4
; Sequence 4, Application US/08354240A
; Patent No. 5670356
; GENERAL INFORMATION:
; APPLICANT: Sherf, Bruce A.
; APPLICANT: Wood, Keith V.
; TITLE OF INVENTION: MODIFIED LUCIFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/354.240A
; FILING DATE: 12-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 34506.029
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-354-240A-4

Query Match 55.8%; Score 1568.5; DB 1; Length 550;
Best Local Similarity 54.7%; Pred. No. 6.1e-152;
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRHGERPRDIVHPGAGOOYQSLYKFASFPE--AIIDAHTNEVSYAQIFETS 57
Db 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDVITYAEYFEMS 60

Qy 58 CRLAVSIOYGLNENNVGVCSENNINFPNVLAAALYLGIPVATNSDMYTDELTHLNI 117
Db 61 VRLAEAMKRYGLNTHRVVCSNSLOQFMPVLGALFICVAVAPANDIYNRELLNSMI 120

Qy 118 SKPTIMESSKKALPLILRVQONLSFIKKVWIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Db 121 SPTVVFVSKKGLQKILNVOKKLPDIIQIIIMDSKTDYQFOSMTFTVTSHPGFNEYD 180

Qy 178 FTKPDFDPLEKIALIMSSGTTGPKGVLSHRSLTIRFVHSRDPYIGTRVTPOTSILSL 237
Db 181 FVPSFDRDKTIALIMSSGTTGPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240

Qy 238 VPFHAFGFMFTLSFVYVGLKVMKPEGALFKTIQNYKIPTIVAPPVNVFLAKSPL 297
Db 241 VPFHGFQFMFTLGLYICGFRVLMYREEEELFRLSLODYKIOSALLVPTLFSFAKSTL 300

Qy 298 VDQYDLSLSEVATGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRFLPCHIROGYGLTETTSAILITPEGDDKPG 360

Qy 358 STGRPLPYIKARVLDNATGKALGPCERGEICFQSEMIKMGYNNPEATIDTDKDGWLHS 417
Db 361 AVGVVPEFEAKVVDLDTGKTLGVNORGLCVRGPMIMSGYNNPEATNALIDKDGWLHS 420

Qy 418 GDIGYDEGDNFFIVDRKELIKYGYVAPAELENLLQHPNIFDAGVAGLPDDAGEL 477
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480

Qy 478 PAACVYLESGLTKEKEVQDFIAAQVTPKHLRGVGFVDSIPKGPCTGLKRLKELREIFA 537
Db 481 PAADVLEHGKMTKEIYDVASQVTTAKKLRGGVGFVDEVPKGLTKLDARKIREILI 540

Qy 538 QRAPKSKL 545
Db 541 KAKRGKGI 548

RESULT 11
US-08-354-240A-6
; Sequence 6, Application US/08354240A
; Patent No. 5670356
; GENERAL INFORMATION:
; APPLICANT: Sherf, Bruce A.
; APPLICANT: Wood, Keith V.
; TITLE OF INVENTION: MODIFIED LUCIFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:

QY 537 -AORAPSKL 545
Db 552 KAKGGRSKL 561

RESULT 13
US-09-380-061B-20
; Sequence 20, Application US/09380061B
; Patent No. 6265177
; GENERAL INFORMATION:
; APPLICANT: SQUIRRELL, DAVID JAMES
; WHITE, PETER JOHN
; LOWE, CHRISTOPHER ROBIN
; MURRAY, JAMES AUGUSTUS HENRY
; TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,061B
; FILING DATE: 25-Aug-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01025
; FILING DATE: 7-APR-1998
; APPLICATION NUMBER: GB 9707468.8
; FILING DATE: 11-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B. J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 124-725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 547 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-380-061B-20

Query Match 54.6%; Score 1534; DB 4; Length 547;
Best Local Similarity 53.5%; Pred. No. 2,1e-148;
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;

QY 1 MEE-ENIRGERPRDIVHPGSAGQOLYOSLYKFAFPE--ALIDAHTNEVSYAQIFETS 57
Db 1 MEDAKNIHMGAPFPYPCDGTAGEDQLHAKRYAQVPTIAFTDHAKEVNIYSEYFEMA 60

QY 58 CRLAVSIEQYGLNENNVYGVCSENNINFPVLAALYIGIPVATSNMYTDBELTGHNI 117
Db 61 CRLAETMRKYGLQHHTAVCSNLSQFMPVCGALFVGCVASTNDIYNEREYNSLSI 120

QY 118 SKPTIMFSKKALPILRVQQLSPKIKVVIDSMYDINGECVSTFVARTDHTFDPLS 177
Db 121 SOPTIVSCSKRALQILGVOKRLPTIQIVILDSREDYMGKQSMYSFTESHLPAFNEYD 180

QY 178 FTPKDPDPLEKIALWSSGGTGLPKGVLSHRSITIRFVHSRDPYIGTRVPTQSI LSL 237
Db 181 YIPDSFDRETALIMSSGGTGLPKGVLSHRSITIRFVHSRDPYIGTRVPTQSI LSL 240

QY 238 VPFHAFGEMFTTSLSYFVVLKVVMLKKFEGALFLKTIQNYKIPTIVVAPPVUMVFLAKSPL 297
Db 241 IPFHGFGEMFTTGLYLCGPRIVLMYRPEBELFRLSLQDYKIOSALLVPTLFSFAKSTL 300

QY 298 VDOYDLSLSEVATGGAPLGGKDAEAVAKRLKPGIIGQYGLTETCCAVMITPHNAVKTG 357
Db 301 VDYDLSNLHEIASGGAPLAKEVGEAVAKRKLPGIRQYGLTETTSALIIITPEGDOKPG 360

QY 358 STGRPLPYIRAKVLONATGKALGPERGEICFQSEMIMKGYNNPEATIDTIDKDGWLHS 417
Db 361 AGKVPVPEFSAKIVDLDTGKTGLGVNQRGELCVKGPMTMKGYNNPEATISALIDKDGWLHS 420

QY 418 GDIGYDEBGNFFIVDRKLKELIKYKGVQVAPAELENLLQHPSTADAGVTPDEFGQL 477
Db 421 GDIAAYDKDGHFFIVDRKLSLIKYGQVPPAELESILLQHPFIFDAGVAGIPDDAGEL 480

QY 478 PAACVWLESKGLTKEKEVQDFIAAOVTPKHLRGVGVFVDSIPKRGPTKLRKELREIFA 537
Db 481 PAAVVLEEGKMTQEQVMDIVAGQVATSKRLRGVGFVDEVPKGLTGKIDGRKIREIL- 539

QY 538 ORAPSKL 545
Db 540 MMGKSKL 547

RESULT 14
US-09-111-752-5
; Sequence 5, Application US/09111752
; Patent No. 6074859
; GENERAL INFORMATION:
; APPLICANT: HIROKAWA, KOZO
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: MURAKAMI, SEIJI
; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,752
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7126-0009-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata and Photinus pyralis
US-09-111-752-5

```
Query Match      54.2%; Score 1524.5; DB 3; Length 552;
Best Local Similarity 54.6%; Pred. No. 2e-147;
Matches 298; Conservative 96; Mismatches 149; Indels 3; Gaps 2;

QY 3 BENIRHGERPRDIVPGSAGQOOLYOSLYKFAFPF-AIIDAHTNEVISYAOIFETSCRLA 61
DB 7 DENIVVGPKEFYPIEGSAGTOLRKYMERAKLGAIAFTNAVTVGDVSYAEYLEKSCCLG 66
QY 62 VSIEOYGLNENNVGVCSENNINFPVLAALYLGIPIVATSDMYTDGELTGHLSKPT 121
DB 67 KALQNYGLVDGRIALSCENCEEFFIPVIAGLFVGVAFTNEIYTLRELVHSLGISKPT 126
QY 122 INFSSKALPLILRQQOQLSFYKVVVVDSDMYDINGECVSTFVARYTDHTDPLSFTPK 181
DB 127 IVFSSKGLDKVITVQKVTITIKTIVILDSKVDYRGYOCLODFIKRNTPPGFQASSFTV 186
QY 182 DPDPLEKIALINSSSGTGLPGVVLSHRSLTIRVHSDPIYGRTRVPTQTSILSLVPPH 241
DB 187 EYDRKEQVALINSSSGTGLPGVQLTHENTVTRFESHARDPIYGNVSPGTAULTVVPFH 246
QY 242 HAFGMFTTLYEVVGLKVVMLKKEGALFLKTIQNYKIPTIVVAPPVVFVLAQSLVDOY 301
DB 247 HGFGMFTTLYLICFRVVMVLTGKDEETFLKTLQDYKCTSVILVPTLIFAILNKSELNKKY 306
QY 302 DLSSLTEVATGAPLGKDVAAEAVAKRLKLPGLIIOGYGLTETCCAVMITPHNAVKTGSTGR 361
DB 307 DLSNLVEIASGAPLSKEVGEAVARRENLPVGRQCYGLTETTTSAIIITPEGDDKPGASCK 366
QY 362 PLPYTKAKVLDNATKALPGERGEICFQSEIMIMKGYNNPEATIDTIDKQWLHSGDIG 421
DB 367 VYPLFKAKVIDLDTKSLGPNRGEVCKVGPMLMKGYNNPEATKELIDEGMLHTGDIG 426
QY 422 YYDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPISIADAGVTGVPDEFGOLPAAC 481
DB 427 YWDEBEHFIVDRLSLKLYKYGVQVAPAELESILLOHPNIFDAGVAGLPDDAGELPAAY 486
QY 482 VYLESGKLTKEVODFIAAQVPTPKHLRGGVVFVDSIPKPGTGLIRKELREIF--AQR 539
DB 487 VYLEHGKMTKEIYDVVASQVTTAKKLKRGVVFVDEVPKGLTGKLDARKIREILIKAKK 546
QY 540 APKSKL 545
DB 547 GGSKL 552
```

```
RESULT 15
US-09-111-752-7
; Sequence 7, Application US/09111752
; Patent No. 6074859
; GENERAL INFORMATION:
; APPLICANT: HIROKAWA, KOZO
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: MURAKAMI, SEIJI
; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,752
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
```

Search completed: September 4, 2002, 15:03:56
Job time: 205 sec